TRUE'S NOTES ON QUANTUM MECHANICS

WILLIAM TRUE
Physics Department
University of California, Davis
Davis, CA 95616, USA

Contents

Some Quantum Mechanics Texts

Schiff — Quantum Mechanics (3rd. Edition) —Text

Messiah — Quantum Mechanics (Vol. I and II)

Davydov — Quantum Mechanics

Baym — Lectures on Quantum Mechanics

Dirac — Quantum Mechanics (4th Edition)

Bohm — Quantum Theory

Merzbacher — Quantum Mechanics

Trigg — Quantum Mechanics

Gottfried — Quantum Mechanics (Vol. I)

Kursunoglu — Modern Quantum Theory

Landau and Lifschitz — Quantum Mechanics

Bethe and Jackiw — Intermediate Quantum Mechanics

Jordan — Linear Operators for Quantum Mechanics

Jauch — Foundations of Quantum Mechanics

Pauling and Wilson — Introduction to Quantum Mechanics

Powell and Crasemann – Quantum Mechanics

Fano — Mathematical Methods of Quantum Mechanics

There are also quite a few quantum mechanics books at the undergraduate level.

1 215A - Introduction and Review

I will assume that you all have had at least one quarter or one semester of undergraduate quantum mechanics. This means that you have been introduced to the wave function $\Psi(\vec{r}, t)$ and the Schrödinger wave equations, $H\Psi = i\hbar\frac{\partial\Psi}{\partial t}$,which tells us how Ψ develops in time. You should have solved the wave equation for the one-dimensional harmonic oscillator, the hydrogen atom, and particles incidents on square potential steps. You probably have also seen or been exposed to many other things.

During this course, we will cover some of the material which you have seen before. But it will be more in the way of a review.

The first part of the course will deal more with various mathematical aspects which relate to quantum mechanics. We will use these things to formulate quantum mechanics in a more general, more powerful, and more useful way than the Schrödinger picture allows.

There are relatively few systems which can be solved exactly. There are even fewer systems in the "real" world which can be solved exactly. So we must resort to approximations in order to describe the physical systems. Much of this course will be devoted to the study of various approsimations and to when and where they can be used.

Later on we will look at some relativistic quantum mechanics and the second quantization picture.

I will not follow any textbook in detail, but much of what I say can be found in Schiff, Messiah, and Boym – and in many other books as well.

I hope the homework will "fill in" some of the gaps in my lectures and give you a better understanding of the methods and techniques used in quantum mechanics.

1.1 $\Psi(\vec{r}, t)$ and $\Phi(\vec{p}, t)$

A Quick Review of Some Points

Usually in a beginning course, one works in the coordinate represenation. We shall see shortly that quantum mechanics can be formulated more generally in a more useful and powerful way by using an abstract vector space known as Hilbert Space. Then our system will be described by a vector in Hilbert Space. As our system evolves in time, this vector will mover around in "our" Hilbert Space.

In the coordinate represenation, the state of a "single" particle system is described by a wave function, $\Psi(\vec{r},t)$, with $|\Psi|^2 d\vec{r}$ being the probability that at time t, the particle will be found in the volume $d\vec{r} = dx\,dy\,dz$ at $\vec{r}$. $\Psi(\vec{r},t)$ is a complex function and must be if $\Psi(\vec{r},0)$ along with the wave equation is to determine $\Psi(\vec{r},t)$ at some later time – including the boundary conditions, of course, (cf. Merzbacher pp. 14-18).

We will generally only consider non-relativistic cases for the first part of the course fo that we do not have to concern ourselves about creation and annihilation of particles and other relativistic effects.

A restriction on $\Psi(\vec{r},t)$ is that is must be "square integrable", i.e., $|\Psi|^2 d\vec{r}$ is a finite real number (It belongs to a Hilbert space). Often, one normalizes Ψ such that $|\Psi|^2 d\vec{r} = 1$, although it is not necessary and, in some cases, not desirable. For example, the plane wave

$$\Psi = Ne^{i(k\cdot r - \omega t)}$$

is not square integrable and couldn't be used as a wave function to describe a moving particle. However, this plane wave can and oftern is used to describe a steady flux of particles.

1.2 $\Psi(\vec{r},t) \rightleftharpoons \Phi(\vec{p},t)$ via Fourier Transforms

Instead of working in the coordinate representation. we could also work in the momentum representation.

In this case, the wave function is $\Phi(\vec{p},t)$ where $|\Phi|^2\, d\vec{p}$ is the probability of finding the particle with momentum $\vec{p}$ in the volume $d\vec{p}$ at time t.

Ψ and Φ are connected to each other and are Fourier transforms of each other. That is,

$$\Phi(\vec{p},t) = \frac{1}{(2\pi\hbar)^{3/2}}\int_0^\infty \Psi(\vec{r},t)e^{-i\vec{p}\cdot\vec{r}/\hbar}\,d\vec{r}$$

and

$$\Psi(\vec{r},t) = \frac{1}{(2\pi\hbar)^{3/2}}\int_0^\infty \Phi(\vec{p},t)e^{-i\vec{p}\cdot\vec{r}/\hbar}\,d\vec{p}$$

Usually, one uses $\vec{p}=\hbar\vec{k}$ and "defines" a $\Phi(\vec{k},t)$ instead of $\Phi(\vec{r},t)$ such that

$$\Phi(\vec{k},t) = \frac{1}{(2\pi)^{3/2}}\int_0^\infty \Psi(\vec{r},t)e^{-i\vec{p}\cdot\vec{r}}\,d\vec{r}$$

and

$$\Psi(\vec{r},t) = \frac{1}{(2\pi)^{3/2}}\int_0^\infty \Phi(\vec{k},t)e^{-i\vec{k}\cdot\vec{r}}\,d\vec{p}$$

Remember that neither Ψ nor Φ can be measurable but only the magnitude of the amplitudes.

How is $|\Psi|^2$ related to the measurement of a particle since a measurement places a particle at a definit point in space? What one does is to "prepare" a large number of identical systems and measure the position of the particle for each system. The measurements will yield a "distribution" of positions and this distribution will approach $|\Psi|^2$ as the number of measurements become very large. Similar remarks can be made concerning the distribution of $|\Phi|^2$.

1.3 The Hamiltonian and the Wave Equation

The Schrödinger wave equation tells us how $\Psi(\vec{r},t$ develops in time. Classically, one has for an isolated system of particles that

$$H(q_1,...q_N,p_1,...p_N,t)=E.$$

The wave equation is given by

$$H\Psi = i\hbar\frac{\partial\Psi}{\partial t}$$

where in H all the p_i's are to be replaced by $\frac{\hbar}{i}\frac{\partial}{\partial q_i}$.

Now one must be careful in following the above prescription. First, H must be written in terms of Cartesian coordinates and their corresponding conjugate momenta. For example, if H was written in terms of spherical polar coordinates and one replaced p_r by $\frac{\hbar}{i}\frac{\partial}{\partial r}$, p_θ by $\frac{\hbar}{i}\frac{\partial}{\partial\theta}$, and p_ϕ by $\frac{\hbar}{i}\frac{\partial}{\partial\phi}$, one

would not obtain the correct wave equation.

Secondly, one must properly symmeterize the combinations of q_i and p_i. For example, pq and qp are the same classically but not quantum mechanically. i.e., $pq\Psi = \frac{\hbar}{i}\frac{\partial}{q}(q\Psi) \neq qp\Psi = q\frac{\hbar}{i}\frac{\partial}{q}\Psi$. In this case, pq must be replaced by the symmetrized expression $\frac{1}{2}(pq + qp)$. To the best of my knowledge, all classical Hamiltonians are such that they can be readily "symmetrized" in the above manner.

We will also encounter systems which have observables which have no classical analogue, e.g., intrinsic spin. In order to write down the Hamiltonian operators, one will have to introduce the operators associated with these "new" observables in a consistent manner. We will discuss this point shortly and only mention here that it is sufficient to give the commutation properties of these new operators with all the other operators of the system.

1.4 Probability Density and Probability Current

We define the "probability density" P as $P = |\Psi(\vec{r},t)|^2 = \Psi^*(\vec{r},t)\Psi(\vec{r},t)$. Its time rate of change is

$$\frac{\partial P}{\partial t} = \Psi^*\frac{\partial \Psi}{\partial t} + \big(\frac{\partial \Psi^*}{\partial t}\big)\Psi$$

and using $H\Psi = i\hbar\frac{\partial \Psi}{\partial t}$, we have with $H = \frac{p^2}{2m} + V(\vec{r})$

$$\frac{\partial P}{\partial t} = -\frac{\hbar}{2mi}\mathrm{div}\big\{\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\big\}.$$

This can be written as

$$\mathrm{div}\vec{S} + \frac{\partial P}{\partial t} = 0$$

by defining the "probability current" $\vec{S}$ as

$$\vec{S} = -\frac{\hbar}{2mi}\big(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\big).$$

$\vec{S}$ describes the "flow" of probability density $P = |\Psi|^2$ just as $\rho\vec{v}$ describes the density flow in the hydrodynamic case by the equation

$$\mathrm{div}(\rho\vec{v}) + \frac{\partial \rho}{\partial t} = 0.$$

Knowing Ψ and Φ we can now write down expressions which tell us what the mean values of

measurements for functions like $F(\vec{r})$ and $G(\vec{p})$ will be. That is

$$\langle F(\vec{r})\rangle = \int |\Psi|^2 \, F(\vec{r}) \, d\vec{r}$$

and

$$\langle G(\vec{p})\rangle = \int |\Phi|^2 \, G(\vec{p}) \, d\vec{p}.$$

Since Ψ and Φ are Fourier transforms of each other which we assume to vanish at infinity, we can always evaluate $F(\vec{r}$ in the momentum representation and/or $G(\vec{p})$ in the coordinate representation. In particular. one can show quite easily that

$$\langle \vec{p}\rangle = \int |\Psi^*(\vec{r},t) \left(\frac{\hbar}{i}\nabla_r\right) \Psi(\vec{r},t) \, d\vec{r}$$

and

$$\langle \vec{r}\rangle = \int |\Phi(\vec{p},t) \left(\frac{\hbar}{i}\nabla_p\right) \Phi(\vec{p},t) \, d\vec{p}.$$

Now all the Ψ's describing a system will form a Hilbert space which we know is a linear space. For example, if Ψ_1 and Ψ_2 belong to this space, then does

$$\lambda_1\Psi_1 + \lambda_2\Psi_2$$

belong to this space where λ_1 and λ_2 are arbitrary complex numbers.

1.5 Scalar Products

In this space we define a "scalar product" as

$$\langle \phi, \psi\rangle \equiv (\phi, \psi) \equiv \int \phi^* \psi \, d\tau$$

where by the last expression I am implying some specific representation, e.g., in the coordinate representation, ϕ and ψ are functions of $\vec{r}$ and $d\tau = d\vec{r}$. The norm will be defined as $\sqrt{\langle \psi, \psi\rangle}$. If $\sqrt{\langle \psi, \psi\rangle} = 1$, the state is said to be normalized.
If $\langle \phi|\psi\rangle = 0$ with $\phi \neq 0$ and $\psi \neq 0$, the two states described by ϕ and ψ are said to be orthonormal.
Some further properties of out scalar product properties are by definition:
1. $\langle \phi|\psi\rangle^* = \langle \psi|\phi\rangle$
2. $\langle \phi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1\langle \phi|\psi_1\rangle + \lambda_2\langle \phi|\psi_2\rangle$

3. $\langle\psi|\psi\rangle \geq 0$ and $\langle\psi|\psi\rangle = 0 \iff \psi = 0$
4. Hermitian conjugate of an operator.
In general, when an operator A operates on a wave function, it changes it into another wave function, e.g., $A\psi = \psi'$. $A^\dagger$ will be defined as the "Hermitian conjugate" of the operator A and has the property that

$$\langle\phi|A\psi\rangle \equiv \langle A^\dagger\phi|\psi\rangle$$

If $A^\dagger = A$, A is said to be a Hermitian Operator. The expectation value of all Hermitian Operators are real. That is

$$\langle A\rangle \equiv \langle\psi, A\psi\rangle = \langle A^\dagger\psi, \psi\rangle = \langle A\psi, \psi\rangle = \langle\psi, A\psi\rangle^*$$

Therefore $\langle A\rangle$ is real.

Using properties 1, 2, and 3 above, one can derive Schwarz's inequality which states that

$$\langle\phi, \phi\rangle\langle\psi, \psi\rangle \geq |\langle\phi, \psi\rangle|^2$$

with the equality sign holding iff $\phi = \lambda\psi$,

There are many operators in quantum mechanics, but all operators corresponding to observables, e.g, position, linear momentum, angular momentum, etc., are Hermitian operators. The measurement of an observable A will generally give a "spread" or "distribution" of values of A. We define the "uncertainty" in A as ΔA where

$$(\Delta A)^2 = \langle A^2\rangle - \langle A\rangle^2.$$

If $\Delta A = 0$, a restriction is placed on the expression $A\psi$ where, of course, A is an Hermitian operator. The expectation value of A is

$$\langle A\rangle = \frac{\langle\psi, A\psi\rangle}{\langle\psi, \psi\rangle}$$

which reduces to $\langle\psi, A\psi\rangle$ if ψ is normalized. If $\Delta A = 0$, we have

$$\frac{\langle\psi, A^2\psi\rangle}{\langle\psi, \psi\rangle} = \frac{\langle\psi, A\psi\rangle^2}{\langle\psi, \psi\rangle^2}$$

and as $\langle\psi, A^2\psi\rangle = \langle A\psi, A\psi\rangle$, we have

$$\langle\psi, \psi\rangle\langle A\psi, A\psi\rangle = \langle\psi, A\psi\rangle^2.$$

This is just Schwarz's inequality with $\phi = A\psi$ and the equality sign holding. Thus $\phi = a\psi$ where a is a complex constant in general. So $A\psi = a\psi$ and then $\langle A \rangle = a$. But A is an Hermitian operator and so $\langle A \rangle$ and therefore a is real.

Using ψ_a instead of ψ above, we have that whenever $\Delta A = 0, A\psi_a = a\psi_a$. This is called an eigenvalue equation with a the "eigenvalue" of the operator A and ψ_a the "eigenfunction".

$e^{ipx/\hbar}$ is an eigenfunction of the operator P_x. i.e.,

$$P_x e^{ipx/\hbar} = \frac{\hbar}{i}\frac{\partial}{\partial x} e^{\imath px/\hbar} = p e^{ipx/\hbar}.$$

But this eigenfunction is not square integrable and does not belong to our Hilbert space. Note that this eigenfunction had a continuous (and not a discrete) spectrum.

Now it is quite easy to show that two eigenfunctions of the operator A with different eigenvalues are orthogonal and linearly independent. To show this, consider

$$\begin{aligned} A\psi_a &= a\psi_a \quad \text{and} \\ A\psi_b &= b\psi_b \end{aligned}$$

Now

$$\langle \psi_b, A\psi_a \rangle - \langle \psi_a, A\psi_b \rangle^* = (a-b)\langle \psi_b, \psi_a \rangle,$$

note that $\langle \psi_b, A\psi_b \rangle^* = \langle \psi_b, A\psi_a \rangle$. Therefore $(a-b)\langle \psi_b, \psi_a \rangle = 0$ and so $\langle \psi_b, \psi_a \rangle = 0$ if $a \neq b$.

To be linearly dependent we need to find non-zero λ_a and λ_b such that $\lambda_a\psi_a + \lambda_b\psi_b = 0$. Taking the scalar product of this last expression with ψ_a, we have

$$\lambda_a \langle \psi_a, \psi_a \rangle + \lambda_b \langle \psi_a, \psi_b \rangle = 0,$$

or $\lambda_a \langle \psi_a, \psi_a \rangle = 0$. But if $\psi_a \neq 0$, we need $\lambda_a = 0$. Likewise, $\lambda_b = 0$ and so ψ_a and ψ_b are linearly

independent.

We see that "non-degenerate" eigenfunctions are orthogonal. If the eigenvalues are equal, it is not clear whether or not they are orthogonal. However, they can be made orthogonal by the "Schmidt orthogonality process" which goes as follows:

Consider the set of eigenfunctions $\psi_1, \psi_2, \psi_3, ..., \psi_N$ all of which have the same eigenvalues, i.e., eigenfunctions are all equal.

1. Take $\phi_1 = c_1\psi_1$ and pick c_1 so that ϕ_1 is normalized, i.e., $c_1^2 = 1/\langle\psi_1, \psi_1\rangle$.
2. Next take $c_2\phi_2 = \psi_2 - \phi_1\langle\phi_1, \psi_2\rangle$.
In this case, we see that $\langle\phi_1, \phi_2\rangle = 0$ and we can pick c_2 such that $\langle\phi_2, \phi_2\rangle = 1$.
3. Next take $c_3\phi_3 = \psi_3 - \phi_1\langle\phi_1, \psi_3\rangle - \phi_2\langle\phi_2, \psi_3\rangle$. We see that $\langle\phi_1, \phi_3\rangle = 0$ and $\langle\phi_2, \phi_3\rangle = 0$ and we pick c_3 such that $\langle\phi_3, \phi_3\rangle = 1$.
4,..,N Just continue on in this manner until one has a set of $\phi_1, \phi_2, ..., \phi_N$ of orthonormal functions where $\langle\phi_i, \phi_j\rangle = \delta_{ij}$.

2 Uncertainty Principle from Schwarz's Inequality

Schwarz'z inequality says that fo two functions f and g

$$(f, f)(g, g) \geq |(f.g)|^2$$

with equality if and only if $f = \lambda g$.

Let us consider two Hermition operators A and B from which we construct two more Hermitian operators $]\alpha = A - \langle A\rangle$ and $\beta - B - \langle B\rangle$. Let $f = \alpha\psi$ and $g = \beta\psi$. Then (

$$(f, f) = (\alpha\psi, \alpha\psi) = (\psi, \alpha^2\psi) = \langle\alpha^2\rangle = \langle A^2\rangle - \langle A\rangle^2 = \Delta A^2$$

and likewise

$$(g, g) = \Delta B^2.$$

Using these in Schwarz's inequality we have

$$\Delta A^2\Delta B^2 \geq |(\alpha\psi, \beta\psi)|^2 = |(\psi, \alpha\beta\psi)|^2 = \frac{1}{4}|(\psi, \{\alpha, \beta\}\psi) + (\psi, [\alpha, \beta]\,\psi)|^2$$

where we used

$$\alpha\beta = \frac{1}{2}(\alpha\beta + \beta\alpha) + \frac{1}{2}(\alpha\beta - \beta\alpha).$$

For Hermitian operators

$$(\psi, \beta\alpha\psi) = (\alpha^\dagger\beta^\dagger\psi, \psi) = (\alpha\beta\psi, \psi) = (\psi, \alpha\beta\psi)^*$$

. Also because α and β are Hermitian we also have the following properties

$$\begin{aligned}(\beta\alpha)^\dagger &= \alpha^\dagger\beta^\dagger = \alpha\beta \\ (\alpha\beta)^\dagger &= \beta^\dagger\alpha^\dagger = \beta\alpha\end{aligned}$$

From which we also have

$$\begin{aligned}\alpha\beta + \beta\alpha &= (\alpha\beta + \beta\alpha)^\dagger \text{ Hermitian: All Real Eigenvalues} \\ \alpha\beta - \beta\alpha &= -(\alpha\beta - \beta\alpha)^\dagger \text{ Anti-Hermitian: All Imaginary Eigenvalues}\end{aligned}$$

Therefore

$$\begin{aligned}(\psi, (\alpha\beta + \beta\alpha)\psi) &= (\psi, (\alpha\beta + \beta\alpha)^\dagger\psi) = 2\Re(\psi, \alpha\beta\psi) \\ (\psi, (\alpha\beta - \beta\alpha)\psi) &= -(\psi, (\alpha\beta - \beta\alpha)^\dagger\psi) = 2\Im(\psi, \alpha\beta\psi)\end{aligned}$$

Therefore the above expression for $\Delta A^2\Delta b^2$ reduces to

$$\Delta A^2\Delta B^2 \geq \frac{1}{4}|(\psi, \{\alpha, \beta\}\psi)|^2 + \frac{1}{4}|(\psi, [\alpha, \beta]\,\psi)|^2.$$

We can always strengthen the inequality by dropping the 1st term on the right hand side. Then

$$\Delta A^2\Delta B^2 \geq \frac{1}{4}|(\psi, [\alpha, \beta]\,\psi)|^2.$$

Let us look at this last expression for the special case where $A = x$ and $B = p$.
Since $[\alpha, \beta] = [A, B] = [x, p_x] = i\hbar$ in this case we have

$$\Delta x\Delta p_x \geq \hbar/2$$

which is the well-known uncertainty principle.
In all cases of two Hermitian operators which commute, we have with $[A, B] = 0$

$$\Delta A^2\Delta B^2 \geq \frac{1}{4}|(\psi, \{\alpha, \beta\}\psi)|^2.$$

Can the equality sign "hold" and if so will $\Delta A\Delta B = 0$? For this to happen, we need $(\psi, \{\alpha, \beta\}\psi) = 0$ and $f = \lambda g$ or $\alpha\psi = \lambda\beta\psi$. This will be the case if ψ is a simultaneous eigenfunction of both A and B.

For example, if $A\psi = a\psi$ and $B\psi = b\psi$, then we know that $\Delta A = \Delta B = 0$. Then $\alpha\psi = \lambda\beta\psi$ with $\lambda = a/b$. Furthermore, it is quite easy to show that $(\psi, \{\alpha, \beta\}\psi) = 0$.
However, in general, ψ need not be a simultaneous eigenfunction of A and B and $(\psi, \{\alpha, \beta\}\psi)$ will not in general be zero. In these cases, we will have the more general relationship $\Delta A \Delta B \rangle 0$ even though A and B commute.

2.1 Special case when $\Delta x \Delta p_x = \hbar/2$

Now let us return to the above one-dimensional case, $\Delta x \Delta p_x \geq \hbar/2$ and inquire what happens when the equality sign holds, i.e., when $\Delta x \Delta p_x = \hbar/2$.
For this to happen, we need $f = \lambda g$ and $(\psi\{\alpha, \beta\}\psi) = 0$. The latter condition tells us

$$0 = (\psi, (\alpha\beta + \beta\alpha)\psi) = (\psi, \alpha g) + (\psi, \beta f) = (\alpha\psi, g) + (\beta\psi, f) = (f, g) + (g, f).$$

Using $f = \lambda g$, we have

$$(\lambda g, g) + (g, \lambda g) = (\lambda^* + \lambda)(g, g) = (\lambda^* + \lambda)\Delta p_x^2 = 0.$$

So for the non-trivial case where $\Delta p_x^2 \neq 0$, we see that $\lambda^* + \lambda = 0$ or that λ is pure imaginary.
Let us now determine $\psi(x)$ for this case. Just to make the math easier, we study the special case where $\langle x \rangle = \langle p \rangle = 0$. Now $f = \lambda g$ becomes $\alpha\psi = \lambda\beta\psi$ or

$$x\psi = \frac{\lambda\hbar}{i}\frac{\partial\psi}{\partial x}.$$

Integrating gives $\psi(x) = N \exp(\frac{ix^2}{2\lambda\hbar})$
We see that for $\psi(x)$ to be zero at $x = \pm\infty$, we need λ to be a negative imaginary number. We already knew it was imaginary.
For convenience, we define $\lambda = \frac{-i}{\nu^2\hbar}$, where ν^2 is a real positive number. Then $\psi = N \exp(\frac{-\nu^2 x^2}{2})$.
Now

$$(\psi, \psi) = 1 = |N|^2 \int e^{-\nu x^2}\, dx = \frac{N^2\sqrt{\pi}}{\nu}$$

and

$$\Delta x^2 = (\psi, x^2\psi) = |N|^2 \int x^2 e^{\nu x^2}\, dx = \frac{|N|^2\sqrt{\pi}}{2\nu^3}$$

allows us to solve for ν^2 and N, i.e.,

$$\nu^2 = \frac{\text{`}}{2\Delta x^2}, \quad \text{and} \quad N = \frac{1}{(2\pi\Delta x^2)^{1/4}}.$$

So for $\Delta x \Delta p = \hbar/2$, we have

$$\psi(x) = \frac{1}{(2\pi\Delta x^2)^{1/4}} \exp^{-x^2/4\Delta x^2}$$

which is a Gaussian shaped wave function centered around $x = 0$ (Because we took $\langle x \rangle = 0$). If $\langle x \rangle$ and $\langle p_x \rangle$ had non-zero values, we would have obtained (cf.Schiff pp 62)

$$\psi(x) = \frac{1}{(2\pi\Delta x^2)^{1/4}} \exp^{\frac{(x-\langle x \rangle)^2}{4\Delta x^2} + i\frac{\langle p_x \rangle x}{\hbar}}.$$

Powell & Grassmann (pp 72) show that if we describe a particle by a Gaussian shaped wave packet then $\Delta x \Delta p_x = \hbar/2$ and we have the maximum information we can obtain about the particle. Furthermore, a Gaussian shaped wave function in the coordinate representation will have a Gaussian shaped wave function in the momentum represenation.
We also see that any other shaped wave function or wave packet will have

$$\Delta x \Delta p_x > \hbar/2.$$

3 The Dirac Delta Function

Consider the relation $y_i = \sum_j a_{ij} x_j$ rewritten as $y(i) = \sum_j a(i,j)x(j)$ which states that $y(i)$ is a linear combination of the $x(i)'s$. If thse indices were continuous, we would have an expression of the form

$$f(x) = \int G(x,x')\, g(x')\, dx'.$$

One says that $f(x)$ is a linear functional of $g(x)$ and $G(x,x')$ is called the "kernel" which depends in general on both x and x'. We see that it is linear because if $f_1 = \int G g_1\, dx$ and $f_2 = \int G g_2\, dx$, then $f_1 + f_2 = \int G(g_1 + g_2)\, dx$. A good example of this is our Fourier transform (in one dimension)

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int e^{ikx} i, \phi(k)\, dk.$$

Here $\psi(x)$ is a linear functional of $\phi(k)$ with a kernel of $\frac{1}{\sqrt{2\pi}} e^{-ikx}$.
On the other hand, we could have considered $y(i) = \sum_j a(i,j)x(j)$ as a linear transformation. Similarly,

$$f(x) = \int G(x,x') g(x')\, dx$$

could be considered as a linear transformation where $g(x)$ are vectors with a set of continuous indices and $G(x,x')$ being the matrix for the transformation.
In the discrete indices case, the "identity" transformation is given by $a(i,j) = \delta_{ij}$. But in the continuous indices case, there is no function of x and x' which does a similar thing. However, this doesn't bother physicists who define a function called the "Dirac delta function", $\delta(x - x')$, which replaces $G(x,x')$ such that

$$f(x) = \int \delta(x - x') f(x')\, dx'.$$

3.1 Common Representations of the δ-function

Some of the more common forms of representation for the Dirac delta function are:

$$\begin{aligned} 1.\ \delta(x) &= \lim_{\epsilon\to 0} \frac{\epsilon}{\pi(\epsilon^2+x^2)} \\ 2.\ \delta(x) &= \lim_{\epsilon\to 0} \frac{1}{\frac{dg}{dx}} \frac{1}{\epsilon(b-a)} \end{aligned}$$

where $g(x)$ is any smooth monotonic function in (a,b) with $g(a)=-\infty$ and $g(b)=+\infty$.

$$\begin{aligned} 3.\ \delta(x) &= \lim_{N\to\infty} \frac{\sin(Nx)}{\pi x} \\ 4.\ \delta(x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{\pm i\mu x}\, du \\ 5.\ \delta(x) &= \lim_{\epsilon\to 0} \frac{e^{-x^2/\epsilon}}{\sqrt{\pi\epsilon}} \\ 6.\ \delta(x) &= \lim_{\epsilon\to 0} \frac{\theta(x+\epsilon)-\theta(x-\epsilon)}{2\epsilon} \to \frac{d\theta(x)}{dx} \end{aligned}$$

where

$$\begin{aligned} \theta(x) &= 0 \text{ for } x<0 \\ \theta(x) &= 1 \text{ for } x>0. \end{aligned}$$

Let us look at the firs one, i.e., $\delta(x-x') = \lim_{\epsilon\to 0} D(x-x',\epsilon)$, where $D(x,\epsilon) = \frac{\epsilon}{\pi(\epsilon^2+x^2)}$. $D(x,\epsilon)$ for three values of ϵ are shown in the Figure at the right. We see that as ϵ becomes smaller, D becomes more peaked around $x=0$ and is "practically" zero elsewhere. Now $\int_{-\infty}^{+\infty} D(x,\epsilon)\,dx = 1$ and is independent of ϵ.

Now if $f(x)$ is continuous around $x\approx x'$, then for small ϵ (sharply peaked around 0), we have

$$\int_{-\infty}^{+\infty} f(x')D(x-x')\,dx' \approx f(x)\int D(x-x',\epsilon)\,dx' = f(x).$$

The above is not a proof. However, if $f(x)$ is bounded everywhere, one can show that

$$lim_{\epsilon\to 0}\int_{-\infty}^{+\infty} f(x)D(x,\epsilon)\,dx = f(0).$$

This is just what we want the δ function to do and if it does it, the exact form of $(D(x,\epsilon)$ is not important.

We will assume that a δ function exists such that

$$\int_{-\infty}^{+\infty} f(x')\delta(x-x')\,dx' = f(x).$$

Furthermore, we should observer that the δ function will only have meaning when it appears under the integral sign.
Later on in this section, we will make a brief digression into Riemann-Stieltjes integrals where we shall see that integrals like the above can be handled with more mathematical rigor.

3.2 3-dimensional δ-functions

Extension into 3 dimensions is "straight forward".

$$\delta(\vec{r}-\vec{r'}) = \delta(x-x')\delta(y-y')\delta(z-z')$$

so that

$$\int f(\vec{r'})\delta(\vec{r}-\vec{r'})d\vec{r'} = \int f(x',y',z')\delta(x-x')\delta(y-y')\delta(z-z')\,dx'dy'dz' = f(x,y,z) = f(\vec{r}).$$

From our work with Fourier transforms

$$\begin{aligned}\delta(\vec{r}-\vec{r'}) &= \frac{1}{(2\pi)^3}\int_{-\infty}^{+\infty} e^{\pm i\vec{k}\cdot(\vec{r}-\vec{r'})}\,d\vec{k}\\ \delta(\vec{p}-\vec{p'}) &= \frac{1}{(2\pi)^3}\int_{-\infty}^{+\infty} e^{\pm i\vec{r}\cdot(\vec{p}-\vec{p'})}\,d\vec{r}.\end{aligned}$$

In spherical coordinates $(r.\theta,\phi)$,

$$\delta(\vec{r}-\vec{r'}) = \frac{\delta(r-r')}{r^2}\sum_{l,m} Y^*_{lm}(\theta.\phi)Y_{lm}(\theta'.\phi'),$$

where the $Y_{lm}(\theta,\phi)$'s are the spherical harmonics.
Example: Let us use the 3rd representation above and show that

$$\begin{aligned}\int_A^B f(x)\delta(x)\,dx &= f(0) \text{ if } A\le 0\le B\\ &= 0 \text{ otherwise.}\end{aligned}$$

We have

$$\int_A^B f(x)\lim_{N\to\infty}\frac{\sin(Nx)}{\pi x}\,dx.$$

We not for x not near zero and N large, the $\sin(Nx)$ oscillates rapidly. Consequently we wouldn't expect a large contribution to the integral from these regions as long as $f(x)$ is a "smooth" and "reasonably well behaved" function.
As $\lim_{x\to 0}\frac{\sin(Nx)}{x} = N$, we will have around $x=0$ just $\int_{x\approx 0} Nf(x)\,dx$. Thus our "main" contributions

should come around $x \approx 0$ and it shouldn't be too surprising to get just $f(0)$.
Let us show this in detal. We let

$$I = \frac{1}{\pi} \lim_{N\to\infty} \int_A^B f(x) \frac{\sin(Nx)}{x}\, dx.$$

Integrate by parts with $u = f(x)$ and $dv = \dfrac{\sin(Nx)}{x}\, dx$.

$$I = \frac{1}{\pi} \lim_{N\to\infty} \left\{ \left[f(x) \int_A^x \frac{\sin(Ny)}{y}\, dy \right]_A^B - \int_A^B f'(x)\, dx \int_A^x \frac{\sin(Ny)}{y}\, dy \right\}.$$

WIth $z = Ny$,

$$I = \frac{1}{\pi} \lim_{N\to\infty} \left\{ \left[f(x) \int_{NA}^{Nx} \frac{\sin(Ny)}{y}\, dy \right]_A^B - \int_A^B f'(x)\, dx \int_{NA}^{Nx} \frac{\sin(Ny)}{y}\, dy \right\}.$$

First, we look at the 1st term on the right hand side. For the lower limit, $Nx \to NA$ and the integral is zero. For the upper limit $Nx \to NB$. If A and B are both $\left\{\begin{matrix}\text{positive}\\ \text{negative}\end{matrix}\right\}$, the integral is still zero as $\lim\limits_{N=\infty} NA = \lim\limits_{N\to\infty} NB = \{\pm\infty\}$. If A is negative and B us positive,

$$\lim_{N\to\infty} \int_{NA}^{NB} \frac{\sin(z)}{z}\, dz = \int_{-\infty}^{+\infty} \frac{\sin(z)}{z}\, dz = \pi.$$

So the 1st term $= f(B)$ if $A < 0$ and $B > 0$ and is zero otherwise.
Similar reasoning applies to the 2nd term. It will be zero unless A is negative and x is positive in the integral over dy. This means that in the integral over dx, we can "neglect" that part it where $x < 0$. So we replace $\int_A^B dx$ by $\int_0^B dx$. With this, the 2nd term becomes

$$-\frac{1}{\pi} \int_A^B f'(x)\, dx \left[\lim_{N\to\infty} \int_{NA}^{Nx} \frac{\sin(Ny)}{y}\, dy \right].$$

The bracket gives π for all $x > 0$ and the second term becomes $-f(B) + f(0)$. So

$$\begin{aligned} I = \int_A^B f(x)\delta(x)\, dx &= f(0) \text{ if } A < 0 \text{ and } B > 0 \\ &= 0 \text{ otherwise} \end{aligned}$$

End example.

3.3 Relations involving the δ-Function

Some of the more "useful" relationships involving the δ function are: (see Schiff pp 57)

1. $\int \delta(x)\,dx = 1$
2. $\int \delta(-x)\,dx = \delta(x)$

   $$i.e. \int_{-\infty}^{+\infty} f(x)\delta(x)\,dx = \int_{-\infty}^{+\infty} f(x)\delta(-x)\,dx = f(0)$$
3. $\delta(ax) = \dfrac{1}{a}\delta(x)$ for $a > 0$
4. $x\delta(x) = 0$

   $$i.e. \int_{-\infty}^{+\infty} f(x)x\delta(x)\,dx = f(x)x\big|_{x=0} = 0$$
5. $\int \delta(x - x'')\delta(x'' - x')\,dx'' = \delta(x - x')$
6. $f(x)\delta(x - x') = f(x')\delta(x - x')$
7. $\delta(x^2 - a^2) = \dfrac{1}{2a}\left[\delta(x - a) + \delta(x + a)\right]$
8. $x\delta'(x) = x\dfrac{d}{dx}\delta(x) = -\delta(x)$
9. $\delta^{(m)}(x) = (-1)^m\delta^{(m)}(-x)$
10. $\int \delta^{(m)}(x - y)\delta^{(n)}(y - a)\,dy = \delta^{(m+n)}(x - a)$
11. $x^{m+1}\delta^{(m)}(x) = 0$
12. $\int f(x)\delta^{(m)}(x)\,dx = (-1)^m f^{(m)}(0)$

    providing that $f^{(m)}(0)$ exists.

Now let us digress briefly and look at Riemann-Stieltjes Integrals. We shall see that the Dirac δ function can be formulated in a "rigorous" manner in this case.

3.4 Digression on Riemann-Steltjes Integrals

(cf. Chapter 9 of Mathematical Analysis by Apostol.)
Riemann Integrals
Many physicists and scientist never see nor never use any integral besides the Riemann integral. For a simplified picture of this integral, let us divide the interval $[a, b]$ up into n parts, let t_k be a point between x_{k-1} and x_k, and let $\Delta x_k = x_k - x_{k-1}$. Then we normally think of the Riemann integral as

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{k=1}^{n} f(t_k)\Delta x_k$$

where we have assumed that the limit exists.
Graphically, this integral represents the area under the curve $f(x)$. Note that $f(x)$ must be continuous but that it does not need to have a continuous first derivative for the integral to exist.
Let us now look at the Riemann-Stieltjes integral which will reduce to the Riemann integral in special cases.

3.5 Riemann-Steltjes Integrals

First we make a few definitions and state some theorems without proof. I refer you to Apostol for details.
We will be considering two functions, $f(x)$ and $g(x)$ where both may or may not be continuous or even differentiable functions.

Definition P is called a partition of the interval $[a, b]$ and consists of the finite set of points $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$.

Definition The partition P' of $[a, b]$ is "finer" than P or a "refinement" of P if P is contained in P'. i.e., $P \subset P'$.

Definition Let $P = \{x_o, x_1, x_2, \ldots x_n\}$ be a partition of $[a, b]$, $\Delta\alpha_k \equiv \alpha(x_k) - \alpha(x_{k-1})$, and t_k be a point in $[x_{k-1}, x_k]$. The sum

$$S(P, f, \alpha) = \sum_{k=1}^{n} f(t_k)\Delta\alpha_k$$

is called a "Riemann-Stieltjes Sum" of f with respect to α on $[a, b]$.

Now we say that f is Riemann integrable with respect to α on $[a, b]$, which we denote by " $f \in R(\alpha)$ on $[a, b]$", if there exists a number A such that:
For every $\epsilon > 0$, there exists a partition P_ϵ of $[a, b]$ such that for every partition P finer than P_ϵ and for every t_k in $[x_{k-1}, x_k]$ we have

$$|S(P, f, \alpha) - A| < \epsilon.$$

When this is the case, we denote A by the integral $\int_a^b f(x)\, d\alpha(x)$ or simply bu $\int_a^b f\, d\alpha$.
One can now show:
If $f \in R(\alpha)$ and $g \in R(\alpha)$, then for arbitrary constants c_1 and c_2, we have

$$\int_a^b (c_1 f + c_2 g)\, d\alpha = c_1 \int_a^b f\, d\alpha + c_2 \int_a^b g\, d\alpha.$$

Similarly, if $f \in R(\alpha)$ and $f \in R(\beta)$, then

$$\int_a^b f\, d(c_1 f + c_2 g) = c_1 \int_a^b f\, d\alpha + c_2 \int_a^b f\, d\beta.$$

So the integral is "linear" in both f and α.
If $c \in [a, b]$ and $f \in R(\alpha)$, then

$$\int_a^b f\, d\alpha = \int_a^c f\, d\alpha + \int_c^b f\, d\alpha.$$

Definition If $a < b$, we define $\int_b^a f\, d\alpha = -\int_a^b f\, d\alpha$ whenever $\int_a^b f\, d\alpha$ exists. Also we define $\int_a^b f\, d\alpha = 0$.

Theorem 3.1 *If $f \in R(alpha)$ on $[a, b]$, then $\alpha \in R(f)$ on $[a, b]$ and*

$$\int_a^b f(x)\, d\alpha(x) + \int_a^b \alpha(x)\, df(x) = f(b)\alpha(b) - f(a)\alpha(a).$$

This "formula" or "relationship" is known as "the formula for integration by parts" and tells us that if $\int_a^b f\, d\alpha$ exists, then so does $\int_a^b \alpha\, df$.
Note as we go along that everything we say also applies to the Riemann integral. However, these relationships are more powerful as we shall see shortly. In fact, f and α need not be continuous and/or differentiable for the above to hold.

Theorem 3.2 *The Riemann Integral.*
If $f \in R(\alpha)$ on $[a, b]$ and $\alpha(x)$ has a continuous derivative on $[a, b]$, then $d\alpha(x)$ can be replaced by $\alpha'(x)\, dx$. Also, the Riemann integral, $\int_a^b f(x)\alpha'(x)\, dx$, exists. We could define $g(x) \equiv f(x)\alpha'(x)$ and then

$$\int_a^b f(x)\, d\alpha(x) = \int_a^b f(x)\alpha'(x)\, dx = \int_a^b g(x)\, dx$$

and the last expression looks more like the integrals we are accustomed to using.

Back to our Riemann-Stieltjes integral. We note that if $\alpha(x)$ is a constant throughout $[a, b]$, then all $\alpha_k = 0$ and $\int_a^b f\, d\alpha$ exists and is zero.
However, out next theorem will state that if $\alpha(x)$ is constant everywhere except for a jump discontinuity at one point, then $\int_a^b f\, d\alpha$ need not exist, but if it does, it need not be zero.

Theorem 3.3 *Given $a < c < b$ and α defined on $[a, b]$ by $\begin{Bmatrix} \alpha(x) = \alpha(a) \text{ for } a \le x < c \\ \alpha(x) = \alpha(a) \text{ for } c < x \le b \end{Bmatrix}$ Let $f(x)$ and $\alpha(x)$ be defined on $[a, b]$ such that at least one of them is continuous from the left at c and at least one of them is continuous from*

the right at c. Then $f \in R(\alpha)$ on $[a,b]$ and we have

$$\int_a^b f\, d\alpha = f(c)\left[\alpha(c^+) - \alpha(c^-)\right].$$

This theorem tells us that we can do integrals for functions f and α like those shown at the right. This is something we couldn't do with Riemann integrals.
We can have several "combinations".

1. $f(x)$ can be both continuous both from the right and from the left and $\alpha(x)$ discontinuous both from the right and the left.
2. $f(x)$ is continuous from the left and discontinuous from the right. Then $\alpha(x)$ must be continuous from the right and discontinuous from the left.
3. $f(x)$ is continuous from the right and discontinuous from the left. Then $\alpha(x)$ must be continuous from the left and discontinuous from the right.
4. If $\alpha(x)$ is continuous from both the right and the left, then $\alpha(c^+) - \alpha(c^-) = 0$ and the integral is zero irrespective of the $f(x)$ does.
5. One can show that if both f and α are discontinuous from the right or from the left, then the integral does not exist.

Now let use use these results to "formulate" the Dirac δ-function using Riemann-Stieltjes integrals. Let $\alpha(x)$ be the step function $\Theta(x-c)$ where $\left\{\begin{array}{l}\Theta(x) = 0 \text{ for } 0 < x \\ \Theta(x) = \text{`` for } x > 0\end{array}\right\}$ Then for a continuous $f(x)$, we have

$$\int_a^b f(x)\, d\alpha(x) = \int_a^b f(x)\, d\Theta(x-c) = f(c).$$

In this case, $d\alpha(x) = d\Theta(x-c)$ behaves just like $\delta(x-c)\, dx$.
So if we would use Riemann-Stieltjes integrals with $\alpha(x)$ being the step function, we would never have to introduce the Dirac δ-function. However, physicists don't do this. They use the Dirac δ-function and Riemann integrals and usually don't run into troubles as long as the functions $f(x)$ are "well behaved" and "so forth".// But if one should get into difficulties, he would have to "back track" and do a more rigorous treatment.
Mathematicians don't stop here. They go further and talk about Lebesgue integrals where they can use functions which are not only discontinuous but which may only be defined in certain regions and at certain points. Then one must become involved with Lesbesgue measure theory.

4 Wave Packets

(cf. Schiff, Messiah, Merzbacher, Gottfied)
In quantum Mechanics, we describe particles by waves of by combinations of waves called wave packets. These wave packets should become more and more localized as one approaches the "classical mechanical region of validity."
In elementary treatment of scattering, one uses plane waves to describe a steady stream of particles incident on a "target" and then scattering. In more formal treatment of scattering, uses of plane waves result in the appearance of mathematical singularities. But if one uses wave packets, as they should, these difficulties disappear. We will see this later in on the year when we consider formal scattering theory.
Using wave packets to describe particles, the inherent "wave properties" will not allow us to make "too precise" statements about position, momentum, etc. and one is forced to do some "hand waving."

4.1 A Simple Wave Packet

In order to gain some insight into all this, let us study some simple wave packets in one dimension in an isotropic non-absorptive medium. Let us consider a disturbance or a pulse which is localized and moving. We would like to have this "packet" describe a moving particle.
The plane wave, $e^{i(kx-\omega t)}$, will have a phase velocity $v_\phi = \omega/k$, where in general ω is a function of k;. Now we could always analyze thie above wave packet in terms of these plane waves – which we attest to "understand" and "know how" to work with them. And it seems reasonably clear that the more "k-components" needed in a Fourier analysis approach, the more localized the wave packet will be and conversely.
We can gain some insight into all this by considering a packet built "equally" from plane waves with k between $k-\Delta k$ and $k+\Delta k$ where $\Delta k \ll k$. In this case, out wave packet is

$$\Psi(x,t) = \int_{k-\Delta k}^{k+\Delta k} e^{i(k'(x-x_0)-\omega' t)}\, dk'$$

where ω' means that $\omega' := \omega(k')$.
Suppressing momentarily the $\omega' t$ part, we look at

$$\begin{aligned} \psi(x) &= \frac{e^{ik(x-x_0)}}{i(x-x_0)}\left[e^{i\Delta k(x-x_0)} - e^{-i\Delta k(x-x_0)}\right] \\ &= 2\frac{\sin[\Delta k(x-x_0)]}{(x-x_0)} e^{ik(x-x_0)} \end{aligned}$$

The real part of $\psi(x)$ is shown at the right and looks something like a wave packet – if it was moving.
A more general approach would be to multiply out exponential by some sort of weighting function,

$f(k')$, which is large near $k' = k$ and small elsewhere and then integrate of k'. e.g. Figures go here – where we did the 2nd example above.
Now

$$|\psi(x)|^2 = \frac{4\sin^2[\Delta k(x - x_0)]}{(x - x_0)^2}$$

and we see that $\psi|^2$ is small when $(x - x_0) \gg \dfrac{1}{\Delta k}$ where Δk is a "measure" of the spread of the wave packet in k-space (momentum space).

4.2 The Relation $\Delta x \Delta k \approx 1$

We required (cf pp 2-4) that $\Delta k \ll k$. So whenever $(x - x_0) \gg \frac{1}{\Delta k}$, then $(x - x_0) \gg \frac{1}{k} = \frac{\lambda}{2\pi}$ which tells us that x is many wavelengths "out from" x_0. Or stating it all inversely, $\psi|^2$ will be large only when

$$\Delta x \Delta k \approx 1$$

where $\Delta x = x - x_0$.
Now this is not an exact statment and we really can't make one when we are dealing with wave packets. All we can say is that $\psi|^2$ will be large when $\Delta x \Delta k \approx 1$. This implies that if Δk, the spread in momentum space, is small, then the spread in coordinate space, Δx, will be large and conversely. This result is also consistent with our uncertainty principle which was $\Delta x \Delta p \geq \hbar/2$ or $\Delta x \Delta k \geq 1/2$.

4.3 General Wave Packets & Group Velocity

Let us now concoct a traveling wave packet by using the weighting function $f(k') = A(k')e^{i\alpha'}$ where $A(k')$ is real and $\alpha' := \alpha(k')$. We have

$$\Psi(x,t) = \int_{-\infty}^{\infty} e^{i(k'x - \omega' t)}\, dk' = \int_{-\infty}^{\infty} A(k') e^{i(k'x - \omega' t + \alpha')}\, dk'$$

where $A(k')$ is expected to be large only around $k \approx k'$. Note in passing that this is just the Fourier transform of the packet, $\Phi(k',t) = \sqrt{2\pi}\, f(k') e^{e} - i\omega' t$ in the momentum representation.

We rewrite $\Psi(x,t)$ as

$$\Psi(x,t) = \int A(k') e^{i\phi'}\, dk'$$

where $\phi' := k'x - \omega' t + \alpha'$. If $\phi' = \phi(k')$ changes rapidly with k', $|\Psi(x,t)|$ will not be large even though $A(k')$ may be large, i.e., $e^{i\phi'}$ oscillates so rapidly, there will be a great deal of cancelation and very little contribution to the integral. Conversely, $|\Psi(x,t)|$ will be large – and this is where the wave

packet will be in the region where ϕ' is stationalry with respect to k'. That is, the "center" of our wave packet will be where

$$\frac{d\phi(k')}{dk'}\Big|_{k'=k} = \frac{d\phi'}{dk'}\Big|_{k'=l} = 0.$$

Let us use the notation, $\frac{d\phi'}{dk'}\Big|_{k'=k} := \frac{d\phi}{dk}, \frac{d\alpha'}{dk'}\Big|_{k'=k} = \frac{d\alpha}{dk}$, etc. Then we have

$$\frac{d\phi}{dk} = x - \frac{d\omega}{dk}t + \frac{d\alpha}{dk} = 0.$$

Since $\frac{d\alpha}{dk}$ is a constant, we see that our wave packet moves to the right with a velocity $v_g := \frac{d\omega}{dk}$ which we call the "group velocity".

4.4 A Wave Packet Striking a Potential Step

Example: As a simple example, let us consider what happens when a wave packet is incident from the right on the potential step shown in the Figure.

Out time independent wave equation is

$$H\psi(x) = \left[\frac{p^2}{2m} + V(x)\right]\psi(x) = E\psi(x).$$

Let $U(x) := \frac{2mV(x)}{\hbar^2}, \mathcal{E} := \frac{2mE}{\hbar^2}$, and using $p^2 = -\hbar^2\frac{d^2}{dx^2}$, our wave equation reduces to

$$\frac{d^2\psi}{dx^2} + (\mathcal{E} - U)\psi = 0.$$

$\mathcal{E}$ can be greater than U_2 or in between U_2 and U_1. We treat these two cases separately.
Case 1: $U_1 < U_2 < \mathcal{E}$ The plane wave solutions in the two regions can be written as (cf. for example, Messiah pp 81)

$$\left.\begin{array}{c}\psi_1 = e^{-ik_1x} + Re^{ik_1x} \\ \psi_2 = Se^{-ik_2x}\end{array}\right\} \text{ where } \begin{cases} k_i = \sqrt{\mathcal{E} - U_i} \\ R = \frac{k_1 - k_2}{k_1 + k_2} \\ S = 1 + R = \frac{2k_1}{k_1 + k_2}\end{cases}$$

These solutions are the ones which would be found for a steady stream of particles incident on the potential step from the right.

In order to study how a wave packet will behave, we will take a linear combination of the ψ's centered around the energy $\mathcal{E}$ – really around k, which is the more relevant parameter in this case. We can construct wave packets by considering a real function $f(k_1' - k_1)$ which is peaked around

$k_1' = k_1$. Then in region 1 we can construct the wave packet

$$\begin{aligned}\Psi_1(x,t) &= \int_0^\infty f(k_1'-k_1)\psi_1 e^{-i\frac{E't}{\hbar}}\,dk_1' \\ &= \int_0^\infty f(k_1'-k_1)e^{-(k_1'x+\frac{E't}{\hbar})}\,dk_1' + \int_0^\infty f(k_1'-k_1)R'e^{i(k_1'x-\frac{E't}{\hbar})}\,dk_1'\end{aligned}$$

and similarly in region 2,

$$\Psi_2(x,t) = \int_0^\infty f(k_1'-k_1)\psi_2(x)e^{-i\frac{E't}{\hbar}}\,dk' = \int_0^\infty f(k_1'-k_1)S'e^{-i)k_2'x+\frac{E't}{\hbar}}\,dk_1'$$

where we have brought in the time dependence in order to have a traveling wave packet.

The integration from 0 to ∞ implies that we have a contribution of all plane waves with $E > V_1$. If $f(k_1'-k_1)$ is sharply peaked, then only those plane waves around $k_1' = k_1$ will give an significant contribution. In the 2nd expression, Ψ_2, we will have to be more restricted on the $f(k_1'-k_1)$ in that we want $E < V_2$. Otherwise we will be including plane wave solutions in our wave packet which do not satisfy the wave equation. So we will assume that we are well above the barrier, i., E well above V_2 and that $f(k_1'-k_1)$ is sufficiently "localized" around k_1 so that the plane wave solutions corresponding to $E < V-2$ to not contribute to the wave packet in regioin 2. To treat the case with $E \geq V_2$ would be more complicated and I don't want to do it here.

Now let us look at the behavior of out constructed wave packets. The 1st term of Ψ_1 is $\int_0^\infty f(k_1'-k_1)e^{-i(k_1'x+\frac{E't}{\hbar})}\,dk_1'$ and represents a wave packet traveling to the left with a center at $x = -v_1 t$ and with a "group" velocity of $v_1 = \dfrac{d\omega}{dk} = +\dfrac{1}{\hbar}\left.\dfrac{dE'}{dk_1'}\right|_{k_1'=k_1}$. (See pages 5 + 6 Section IV where we discussed the velocity of the center of a wave packet).

We have $E - V_1 = \dfrac{p_1^2}{2m} = \dfrac{\hbar^2 k^2}{2m}$ and so $v_1 = \dfrac{\hbar k_1}{m} = \dfrac{p}{m}$. So in this case, the center of the wave packet moves with the classical velocity. Since $x = -v_1 t$, we see that in the remove past, the center of the wave packet was far to the right and moved with a constant velocity to the left reaching the origin at time $t = 0$.

Similarly, the 2nd term in Ψ_1 represents a wave packet moving to the right with the same velocity v_1 with its center at $x = v_1 t$. So this "reflected" packet starts out at the origine at time $t = 0$ and moves with a constant velocity to the right.

We can ask what in the above precludes the possibility that this "reflected" packet was already moving to the right at a large distance in the remote past. i.e., long before out incident packet, the 1st term in Ψ_1, could reach the potential step at $x = 0$ and be reflected. In the 2nd term we have a phase in the exponent of $\phi' = k_1' - \dfrac{E't}{\hbar}$. For large x and $t \ll 0$ (a large negative t to the remote past), ϕ' is a positive quantity with $e^{i(k_1'x-\frac{E't}{\hbar})}$ oscillating rapidly as we integrate over k_1'. Consequently

there will be essentially no contribution to the integral in this case. This Ψ_1(2nd term) will only be "large" when the phase of ϕ' is statioinary i.e., x and t both positive with $x = v_1 t$. Thus the reflected packet does not "appear" until the incident packet arrives and is reflected.

Similar arguments can be made for the incidence packet Ψ_1 (1st term), which has a phase $\phi' = k_1' + \frac{E't}{\hbar}$. This packet will only exist for $x > 0$ and $t < 0$ with $x = -v_1 t$. At $t = 0$, when it reaches the potential step, it "disappears" never to be seen again.

Using the same kind of arguments, Ψ_2 represents a wave packet which "originates" at the origin at time $t = 0$ and travels to the left. As $\phi' = k_2' x + \frac{E't}{\hbar}$, the phase is stationary at

$$\left.\frac{dk_2'}{dk_1'}\right|_{k_1'=k_1'} x + \left.\frac{dE'}{dk_1'}\right|_{k'=k} \frac{t}{\hbar} = \frac{dk_2}{dk_1} x + v_1 t = 0.$$

As $k_1^2 = k_2^2 + U_2 - U_1$, $\frac{dk_2}{dk_1} = \frac{k_1}{k_2} = \frac{v_1}{v_2}$. and we see that the center of our "transmitted" packet is at $x = -v_2 t$ traveling with its classical velocity to the left.

Note that the reflected packet is reduced by R from the incident packet and the transmitted packet is reduced by S. The only difference between this approach and a classical picture is that a classical particle would not have been reflected but would have continued on into region 2 with a decreased velocity of v_2. Light waves on the other hand would be "reflected" and "transmitted" when they struck an interface between two dielectric mediums.

Case 2: $U_1 < \mathcal{E} < U_2$ Now we look at the other case where the energy is below V_2. Things will go much the same here as in Case 1 but with some important differences.

The plane was solutions can be written as (see Messiah pp 82 for example)

$$\left.\begin{array}{rcc} \psi_1 & = & e^{-ik_1 x} - e^{ik_1 x + 2\phi} \\ \psi_2 & = & A e^{i\phi} e^{k_2 x} \\ & & \text{decaying exponential as } x \text{ is negative} \end{array}\right\} \text{ where } \left\{\begin{array}{rcc} k_1 & = & \sqrt{\mathcal{E} - U_1} \\ k_2 & = & \sqrt{U_2 - \mathcal{E}} \\ A & & \text{is a constant} \\ \sin\phi & = & \sqrt{\frac{\mathcal{E}-U_1}{U_2-U_1}} = \frac{k_1}{\sqrt{k_1^2+k_2^2}} \end{array}\right.$$

Using $f(k_1' - k_1)$, we construct the packets

$$\begin{aligned} \Psi_1(x,t) &= \int_0^\infty f(k_1' - k_1) e^{-(k_1' x + \frac{E't}{\hbar})}\, dk_1' + \int_0^\infty f(k_1' - k_1) e^{i(k_1' x + 2\phi' - \frac{E't}{\hbar})}\, dk_1' \quad \text{and} \\ \Psi_2(x,t) &= A e^{k_2 x} \int_0^\infty f(k_1' - k_1) e^{-(k_1' x + \frac{E't}{\hbar})}\, dk_1'' \end{aligned}$$

using a real $f(k_1' - k_1)$ centered around $k_1' = k_1$.

By similar arguments to those given on 8-IV, $f(k_1' - k_1)$ should be zero for k_1''s corresponding to $E > V_2$ and we will stay away from the top of the barrier where $E \leq V_2$.

As in Case 1 above, the 1st term in Ψ_1 represents a wave packet traveling to the left with the classical velocity v_1, centered at $x = -v_1 t$, reaching the origin at $t = 0$, and "disappearing".

Similarly, the 2nd term in Ψ_1 represents a packet, non-existent for $t < 0$, traveling to the right with the same velocity, v_1, which as its center at

$$x = v_1 t - 2\frac{d\phi'}{dk_1'}\bigg|_{k_1'=k_1}.$$

We note that the center of the incident packet reaches the origin at $t = 0$ while the center of the reflected packed doesn't leave the origin until a time $t = \frac{2}{v_1}\frac{d\phi}{dk_1}$ later. now how significant is this "delay" time and can we gain some insight into its meaning? The answer is both Yes! and No!

If we want to compute the motion of our packet with a classical particle, then it should retain a nice localized shape – at least away from the origin where the potential step is. Near the origin we can expect some distortion as the "front part" of the packet arrives first and is reflected before the "back part" of the packet arrives. In order to have a nice localize packet, we don't want the $2\phi'$ term in the phase varying too rapidly with k_1' over the range Δk where $f(k_1' - k_1)$ is large. That is, we want

$$\Delta k \left(2\frac{d\phi}{dk_1}\right) \ll 1.$$

or using $\Delta x \Delta k \approx 1$, we want $\Delta x \gg 2\frac{d\phi}{dk_1} = v_1\tau$. This last relationship tells us that the delay time, τ, is much less than the time it takes the packet to move its mean width Δx. Since we are considering a bundle of waves, it is not reasonable to ask details about what goes on during such a short time.

But we can gain some insight into what is going on by looking at $\Psi_1(x,t)|$. $|\Psi_2|$ will only be large when $\left[\frac{d}{dk_1;}(\phi' - \frac{E't}{\hbar})\right] \approx 0$. That is, when $\frac{d\phi}{dk_1} = \frac{t}{\hbar}\frac{dE}{dk_1}$ which is when $t = \tau/2$. In constrast to a classical particle which never gets into region 2. the packet has a good chance of existing in region 2 near the origin for a time of order $\tau/2$. How far it gets into region 2 depends on the size of k_2. So this is purely a quantum mechanical result where the particle penetrates into region 2, finds out that it "shouldn't be there" and "gets out".

Could we ever detect the particle in there? One can also ask similar question such as can a baseball go through a glass window without breaking it?

4.5 Spreading of the Wave Packets

Let us now see how wave packets in general will spread. Qualitatively we expect this to happen as a packet is made up of several momenta and one would expect the faster components to "get ahead" of the slower ones.

Now a particle in a potential which is independent of its momentum, viz.

$$E = \hbar\omega = \frac{p^2}{2m} + v(x),$$

has a group velocity of $v_g = \dfrac{\partial\omega}{\partial k} = \dfrac{\hbar k}{m} = \dfrac{p}{m} = v_c l$. we see that the center of this packet moves with a constant velocity, the classical velocity, $v_c l$, but the packet will still spread as we shall see below.

Dropping the prims in the integration, we consider a packet with a Gaussian shaped weighting function, $f(k-k_0) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{(k-k_0)^2}{2\Delta k^2}}$, which is centered around $k = k_0$ with a width Δk defined by $\Delta k^2 = \left[\langle k^2\rangle - \langle k\rangle^2\right]_{t=0}$. Furthermore, let's assume that the packet is centered at x_0 at time t_0.

Then our packet is given by

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(k-k_0)^2}{2\Delta k^2}}\, e^{i[k(x-x_0)-\omega t]}\, dk.$$

Now we can't do the integration until we know how ω varys with k. If it was a free particle, $E = \hbar\omega = \dfrac{p^2}{2m} = \dfrac{\hbar^2 k^2}{2m}$ or if the particle moved in a momentum independent potential, $E = \hbar\omega = \dfrac{\hbar^2 k^2}{2m} - V(x)$, then we would know how ω depended on k and we can go ahead an integrate.

Even if we didn't know how ω depended on k but if Δk was small enough, we could hope to expand ω about $k = k_0$ and neglect the higher order terms as being small. That is, the higher order terms will be small for small $k - k_0$ and will be "cut off" by $f(k-k_0)$ for large $(k-k_0$. so

$$\omega(k) = \omega(k_0) + (k-k_0)\left.\frac{\partial\omega}{\partial k}\right|_{k=k_0} + \frac{1}{2}(k-k_0)^2 \left.\frac{\partial^2\omega}{\partial k^2}\right|_{k=k_0} + \ldots.$$

Let use just keep the 1st three terms and define $K = (k-k_0), X = x - x_0, \omega_0 = \omega(k_0), v_g = \left.\dfrac{\partial\omega}{\partial k^2}\right|_{k=k_0}$ and $\alpha = \left.\dfrac{\partial^2\omega}{\partial k^2}\right|_{k=k_0}$.

With all this, our wave packet become

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{K^2}{2\Delta k^2}} e^{i\left[Kx - v_g K t - \frac{1}{2}\alpha K^2 t\right]} e^{i[k_0 X - \omega 4]}\, dK.$$

We can complete the square in the exponents of the 1st two terms obtaining an integral of the form

$$\int_{-\infty}^{\infty} e^{-a^2 y^2}\, dy = \frac{\sqrt{\pi}}{a}.$$

This procedure give us

$$\Psi(x,t) = \sqrt{\frac{\Delta k^2}{1+i\alpha\Delta k^2 t}}\, e^{-\left[\frac{(x-v_g t)^2\Delta k^2}{2(1+\Delta k^2\alpha^2 y^2}\right]}\, e^{i\left[\frac{\alpha\Delta k^4+(x-v_g t)^2}{2(1+\alpha^2\Delta k^4 t^2}+k+0x=\omega_0 t\right]}$$

from which

$$|\Psi(x,t)|^2 = \left|\frac{\Delta k^2}{(1+i\alpha\Delta k^2 t)}\right| e^{-\left[\frac{\Delta k^2(x-v_g t)^2}{1+\alpha^2\Delta k^2 t^2}\right]}.$$

If $\alpha = 0$, we see that we will have a wave packet of constant shape centered around $x = x_0 + v_g t$. But if $\alpha \neq 0$, we see that the amplitude of the packet will decrease in time and that the packet will also "spread" in time. A measure of this spreading can be taken as the width of the packet at $1/e$ of its maximum value, i.e.,

$$\Delta x \approx \sqrt{\frac{1+\alpha^2\Delta k^t t^2}{\Delta k}}.$$

Let $\delta x|_{t=0} = \Delta x \approx \dfrac{1}{\Delta k}$. Then we have

$$\delta x = \Delta x\sqrt{1+\frac{\alpha^2 t^2}{\Delta x^4}}.$$

We see for small times, where $\dfrac{\alpha^2 t^2}{\Delta x^4} \ll 1$, that $\delta x \approx \Delta x$ and the packet hasn't spread much. For much longer times, where $\dfrac{\alpha^2 t^2}{\Delta x^4} \gg 1$, we have $\delta x = \dfrac{\alpha t}{\Delta x}$ and the spreading increases linearly with t.

For a free particle, $\left(E = \dfrac{p^2}{2m}\right)$, or for a particle moving in a potential, $E = \dfrac{p^2}{2m} + V(x)$, we have $\alpha = \left.\dfrac{\partial\omega}{\partial k}\right|_{k=k_0} = \dfrac{\hbar}{m}$. So all packets of interest will spread as they move along. For classical particles, it doesn't affect us because $\dfrac{\hbar}{m}$ is very small.

5 Mathematical Framework of Quantum Mechanics

Much of what I discuss here will be found in Chapter VII of Messiah. He prbably does the most complete and detailed treatment of the pertinent mathematics which relates to quantum mechanics. He is, in fact, more detailed that I will be here and I urge you to read this Chapter.

Most beginning graduate students, and most likely you are no exceptions, have worked with quantum mechanics in the "coordinate represenation" or the "Schröfinger Picture" as it is commonly called. But one could also work in the "momentum representation" where $\vec{p}$ is replaced by the operator $\vec{p}$ and $\vec{x}$ replaced by the operator $-\frac{\hbar}{i}\nabla_{\vec{p}}$. In fact, if you write down the differential equation to be solved in the momentum representation for the harmonic oscillator, you will find that you have the

same differential operator to solve as you did in the "coordinate represenation." Of course, constant coefficients may be changed in the differential equation in the two representations. Then one can ask the question: – Are there other representations which one could work in? The answer to this is "Yes"! In fact, there are in infinite number of them. In fact, one can solve the wave equation without ever going into a representation.–A little later, I will show this explicitly where I will solve the harmonic oscillator without going into any representation. Often times in more complicated systems, it "complicates things" to be in a specific representation.

Now we will want to formulate Quantum Mechanics in a more abstract way so that it is independent of the representation. We will then show that if we go into the coordintate representation, it will reduce to out "familiar" Schrödinger picture. Let us now define a field.

5.1 Definition of a Field

A set of scalars $\{\alpha, \beta, \gamma, ...\}$ form a field, F, if the scalars have the following properties:

A. To every pair, α and β, there corresponds a scalar in the field, $\alpha + \beta$, called the sum of α and β in such a way that:

1. $\alpha + \beta = \beta + \alpha$ — Communicative
2. $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ — Associative
3. $\alpha + 0 = \alpha$ — Null Scalar Exists
4. $\alpha + (-\alpha) = 0$ — Inverse Exists

and B. To every pair, α and β, there corresponds a scalar, $\alpha\beta$, in such a way that

1. $\alpha\beta = \beta\alpha$ — Communicative
2. $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ — Associative
3. $\alpha I = \alpha$ — A unique unit scalar exists
4. $\alpha\alpha^{-1} = I$ for$\alpha \neq 0$ — An inverse exists

and C.

$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ — Multiplication is distributive with respect to addition.

Some examples of a field under regular addition and multiplication:

1. The set of all real rational numbers, $\{Q\}$.
2. The set of all real numbers, $\{R\}$.
3. The set of all complex numbers, $\{C\}$.

We now define a linear vector space.

5.2 Definition of a Linear Vector Space

Consider the scalars, α, β, γ,..., of the field F. Also consider a set of elements. $|x\rangle$, $|y\rangle$, $|z\rangle$,... called vectors. This set of vectors form a linear vector space, V, when they satisfy:

A.) To every pair, $|x\rangle$ and $|y\rangle$, there corresponds a vector $|x\rangle + |y\rangle$ in V called the sum of $|x\rangle$ and $|y\rangle$ in such a way that:

1. $|x\rangle + |y\rangle = |y\rangle + |x\rangle$ — Communicative
2. $|x\rangle + (|y\rangle + |z\rangle) = (|x\rangle + |y\rangle) + |z\rangle$ — Associative
3. $0 = |0\rangle$ and $0 + |x\rangle = |x\rangle + 0 = |x\rangle$ — A null vector exists
4. $|x\rangle + (-|x\rangle) = 0$ — an inverse exitst.

B.) To every pair α and β in F and $|x\rangle$ in V, there corresponds a vector $\alpha|x\rangle$ in V, called the product of α and $|x\rangle$, such that:

1. $\alpha(\beta|x\rangle) = (\alpha\beta)|x\rangle$ — Associative
2. $I|x\rangle = |x\rangle$

C.) 1. $\alpha(|x\rangle + |y\rangle) = \alpha|x\rangle + \alpha|y\rangle$ — Multiplication by a scalar is distributive with respect to vector addition

2. $(\alpha+\beta)|x\rangle = \alpha|x\rangle + \beta|y\rangle$ — Multiplication by a vector is distributive with respect to scalar addition.

We continue on with more mathematical definitions and theorems (which I won't always prove).

Definition of Span

The set of vectors, $\{|x_1\rangle, |x_2\rangle, ..., |x_n\rangle\}$ are said to Span, or Generate, the space V if any vector, $|x\rangle \in V$ is expressible as a linear combination of them, i.e., if $|x\rangle = \lambda_1|x_1\rangle + \lambda_2|x_2\rangle + ... + \lambda_n|x_n\rangle$ for some scalars $\lambda_1, \lambda_2, ..., \lambda_n \in F$, which are not necessarily unique.

Definition of Linearly Independent

In the abstract vector space V the finite set of vectors $\{|x_1\rangle, |x_2\rangle, ..., |x_n\rangle\}$ is said to be linearly dependent if scalars $\lambda_1, \lambda_2, ..., \lambda_n \in F$ exist, not all zero, such that $|x\rangle = \lambda_1|x_1\rangle + \lambda_2|x_2\rangle + ... + \lambda_n|x_n\rangle = |0\rangle$. If no such scalars exist, i.e., if $|x\rangle = \lambda_1|x_1\rangle + \lambda_2|x_2\rangle + ... + \lambda_n|x_n\rangle = |0\rangle \Rightarrow \lambda_i = 0 \forall i = 1, ..., n$ then the set is linearly independent.

Definition of a Basis

The set $S = \{|x_1\rangle, |x_2\rangle, ..., |x_n\rangle\}$ is a basis of V if (i) they are linearly independent and (ii) they span V. If V possesses a (finite) basis it is said to be finite dimensional, if not then it is infinite dimensional.

Theorem 1 *If $S = \{|x_1\rangle, |x_2\rangle, ..., |x_n\rangle\}$ span V then they form a basis if and only if any vector $|x\rangle$ in V is uniquely expressible as a linear combination of the elements of S.*

Theorem 2 *If $S = \{|x_1\rangle, |x_2\rangle, ..., |x_n\rangle\}$ span V then there is a subset of these which is a basis of V.*

Theorem 3 *If* $S = \{|x_1\rangle, |x_2\rangle, ..., |x_p\rangle\}$ *are linearly independent and* V *is finite dimensional then there exists a base containing* $\{|x_1\rangle, |x_2\rangle, ..., |x_p\rangle\}$.

Theorem 4 *If* $\{|x_1\rangle, |x_2\rangle, ..., |x_p\rangle\}$ *is a linearly independent set and* $\{|y_1\rangle, |y_2\rangle, ..., |y_m\rangle\}$ *spans* V*, then* $p \leq m$,

Theorem 5 *The number of elements in any basis of a finite-dimensional vector space is the same as in any other basis.*

The proof of this case can go as follows:
Consider two sets of vectors, $S_1 = \{|x_1\rangle, |x_2\rangle, ..., |x_n\rangle\}$ and $S_2 = \{|y_1\rangle, |y_2\rangle, ..., |y_m\rangle\}$ where S_1 spans the space but all the elements in S_1 may or may not be linearly independent, and all the elements in S_2 are linearly independent but S_2 may of may not span the space.

If S_2 does not span the space, we can find an element in S_1, callit $|x_1'\rangle$, which is linearly independent of the $|y_i\rangle$'s. If $S_2' = \{|y_1\rangle, |y_2\rangle, ..., |y_m\rangle, |x'\rangle\}$ does not span the space, we can find another linearly independent vector $|x_2'\rangle$ from S_1 and add it to S_2' and so on and so forth until $S_2^p = \{|y_1\rangle, |y_2\rangle, ..., |y_m\rangle, |x'\rangle, ..., |x_p'\rangle\}$ spans the space and forms a basis. Clearly $p < n$ and $m \leq n$. Reversing the rols of S_2 and S_2, we would find $m \geq n$ and so $m = n$ when S_1 and S_2 are both bases.

Definition of Dimensions
The dimension of a finite dimensional vector space is the number of elements in a basis.

Theorem 6 *Every* $n + 1$ *vectors in an* n*-dimensional vector space is are linearly dependent.*

Definition of Isomorphism Two vector spaces U and V over the same field F are said to be isomorphic to each other if there is a one-to-one correspondence between the vectors $|x\rangle \in U$ and the vector

$|y\rangle \in V$, say $|y\rangle = T(|x\rangle)$, such that

$$T(\alpha_1|x_1\rangle + \alpha_2|x_2\rangle) = \alpha_1 T(|x_1\rangle) + \alpha_2 T(|x_2\rangle).$$

That is, all linear relationships are preserved.

Theorem 7 *Any finite dimensional vector space V is isomorphic to the space of n-dimensional co-ordinate vectors, the co-ordinates being members of the scalar field F and n being the dimension of V.*

5.3 Definition of an Unitary Space

A "Unitary Space" is a linear vector space such that for any two vectors, $|x\rangle$ and $|y\rangle$, a unique scalar – called the "inner product" or "scalar product" and denoted $\langle x|y\rangle$ exists and has the following 4 properties:

1. $\langle x|y\rangle = \langle y|x\rangle *$
2. $\langle x|y+z\rangle = \langle x|y\rangle + \langle x|z\rangle$
3. $\langle x|\alpha y\rangle = \alpha\langle x|y\rangle$
4. $\langle x|x\rangle \geq 0$ where $\langle x|x\rangle = 0 \iff |x\rangle = 0$.

This scalar product is essentially the same as the one we introduced in Section I except it "applies" to vectors instead of scalar functions.

When dealing with finite dimensional unitary spaces, one doesn't have too much trouble. However, for infinite dimensional unitary spaces which we often have to deal with in quantum mechanics, we can have "infinities" appearing unless we place a further restrictions on this space. Consequently, we will only consider Unitary Spaces which are complete.

5.4 Definition of a Hilbert Space

A complete Unitary Space is called a Hilbert Space. Equivalently, a complete linear vector space of finite of infinite dimension with scalar product is a Hilbert Space.

By being complete, we mean that for a sequence of vectors, $|x_n\rangle$, a vector $|x\rangle$ exists such that

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

I realize that I am not being complete and perhaps rigourous here. But I don't want to take the time here and refer you to mathmatical texts dealing with Linear Vector Spaces, Unitary Spaces, and Hilbert Spaces.

Example of a Hilbert Space

Consider the set of all infinite sequencees, $\{[x_i], [x_i'], ...\}$, wherer $[x_i]$ is an infinite sequence and $\sum_i \|x_i\|^2$ is finite. This set of sequences form a Hilbert Space.

2nd Example of a Hilbert Space

Consider the set of real variables, $q_1, q_2, ..., q_k$, defined over some regions and the set of all functions $f(q_1, q_2, ..., q_k)$ such that

$$\int\int ... \int |f(q_1, q_2, ..., q_k)|^2 \, dq_1 dq_2 ... dq_k < \infty.$$

This set of functions form a Hilbert Space where the scalar product for two of the functions, f and g, is defined as

$$\langle f|g\rangle = \int\int ... \int f * g \, dq_1 dq_2 ... dq_k.$$

In fact, this is just the space of square-integrable functions which we use the the Schrödinger picture.

Where are we going? And why all this mathematics? We shall see in the next secton that one of our postulates of Quantum Mechanics will be:
To every type of physical system, there will correspond an abstract Hilbert Space and each vector in this space will represent exactly one possible state of the sytem.

5.5 Linear Operators

Let use consider the expression

$$|x\rangle = A|y\rangle.$$

By this, we mean that an operator A operators on the vector $|y\rangle$ and "produces" a new vector $|x\rangle$. A simple example wold be the operator which rotates a vector around some axis in 3-dimensional space.

We will not be interested in operators which take us outside our Hilbert spaces. We can and occasionally do run into operators which take use from one Hilbert Space to another. For example, Let $HS-1$ be the space of square integrable functions $f(x_3)$ and HS_3 be the space of square integrable functions $F(x_1, x_2, x_3)$. The linear operator $Q = \alpha \int dx_1, dx_2$ will give us $QF = f$ and we see that Q takes us from HS_3 to HS_1.

Usually we will only be interested and will deal with linear operators although we will occasionally run into antilinear operators. An operator A is said to be "linear" if

$$A(|x\rangle + |y\rangle) = A|x\rangle + A|y\rangle \text{ and } A(\alpha|x\rangle) = \alpha A(|x\rangle)$$

or

$$A(\alpha|x\rangle + \beta|y\rangle) = \alpha(A|x\rangle) + \beta(A|y\rangle$$

where α and β are arbitrary complex numbers.

An operator A is said to be "anti-linear" if

$$A(\alpha|x\rangle + \beta|y\rangle) = \alpha^*(A|x\rangle) + \beta^*(A|y\rangle).$$

Two operators A and B are said to be equal if

$$A|x\rangle = B|x\rangle \text{ for all } |x\rangle.$$

also

$$\begin{aligned} (A+B)|x\rangle &= A|x\rangle + B|x\rangle, \\ (AB)|x\rangle &= A(B|x\rangle), \text{ and} \\ AB &\neq BA \text{ in general.} \end{aligned}$$

5.6 Dirac's Bra and Ket Notation 1

In the foregoing, we were careful to designate our vectors by the symbols $|x\rangle, |y\rangle, \ldots$ and our scalar products by $\langle x|y\rangle, \langle x|z\rangle, \ldots$. Dirac calls these vectors $|x\rangle, |y\rangle, |z\rangle$ by the name of "kets". Furthermore, he defines $\langle y|$ as a "bra" which is conjugate to the ket $|y\rangle$. Then $\langle y|x\rangle$ becomes a "bra-ket".

Now to every ket $|x\rangle$ in our vector space, we will have a unique bra $\langle x|$. Since the $|x\rangle$'s belong to a linear space, it is not hard to see that the bras $\langle x|$ will also form a linear vector space. The linear vector space of the $\langle x|$'s is called the space dual to the linear space of the $|x\rangle$'s.

At this point one can or cannot use explicitly this concept of dual space. Both Messiah and Dirac do and it does make things more "clearer" when one deals with anit-linear operators, etc. For this reason, I will digress and discuss this dual space which consists of all the linear functionals of the $|x\rangle$'s.

Before we can do that let us note that the property of our scalar product

$$\langle x|z\rangle = \langle z|x\rangle^*$$

requires that the bra conjugate to the ket

$$|x\rangle = \alpha_1|y_1\rangle + \alpha_2|y_2\rangle$$

is

$$\langle x| = \alpha_1^*\langle y_1| + \alpha_2^*\langle y_2|.$$

that is, the relationship is anti-linear. That this must be the case is clear since from the above

$$\begin{aligned} \langle z|x\rangle &= \alpha_1\langle z|y_1\rangle + \alpha_2\langle z|y_2\rangle \text{ and} \\ \langle x|z\rangle &= \alpha_1^*\langle y_1|z\rangle + \alpha_2^*\langle y_2|z\rangle \end{aligned}$$

Now we digress and look at some of the aspects of our dual space.

5.7 The Dual Space 1

Consider (in obvious shorthand notation) the vector space V with vectors $x_1, x_2, x_3, \ldots$

Definition A linear functions on a vector space V is a scalar-valued function $y(x)$ defined such that for every vectr $x = \alpha_1 x_1 + \alpha_2 x_2$ we have

$$y(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y(x_1) + \alpha_2 y(x_2)$$

where α_1 and α_2 are scalars from the field F.

An example of a linear functional, consider

$$y_1(x_1) = \int A_1(t)x_1(t)\,dt.$$

In this case, we say y_1 is a linear functional of x_1 on V. Clearly we can have $y_1(x_i)$ where $i = 1, 2, 3, \ldots$. Now consider another linear functional $y_2(x_i)$ where

$$y_2(x_i) = \int A_2(t)x_i(t)\,dt.$$

Then every $y(x) = \alpha_1 y_1(x) + \alpha_2 y_2(x)$ is also a linear functional if y_1 and y_2 are linear functionals. For our simple example, this last statement is obvious. Let us now show it in general.

What we want to show is that if $y_1(x)$ and $y_2(x)$ are linear functionals, then so is $y(x) = \alpha_1 y_1(x) + \alpha_2 y_2(x)$. To do this consider $x = \beta_1 x_1 + \beta_2 x_2$. Then

$$y(x) = y(\beta_1 x_2 + \beta_2 x_2) = \alpha_1 y_1(\beta_1 x_1 + \beta_2 x_2) + \alpha_2 y_2(\beta_1 x_1 + \beta_2 x_2).$$

Since y_1 and y_2 are linear functionals, we have

$$\begin{aligned} y(x) = y(\beta_1 x_1 + \beta_2 x_2) &= \alpha_1 \left[\beta_1 y_1(x_1) + \beta_2 y_1(x_2)\right] + \alpha_2 \left[\beta_1 y_2(x_1) + \beta_2 y_2(x_2)\right] \\ &= \beta_1 \left[\alpha_1 y_1(x_1) + \alpha_2 y_1(x_1)\right] + \beta_2 \left[\alpha_1 y_2(x_2) + \alpha_2 y_2(x_2)\right] \\ &= \beta_1 y(x_1) + \beta_2 y(x_2) \end{aligned}$$

So we see that $y(x)$ is also a linear functional. Consequently, any linear combination of linear functionals is also a linear functional.

Let the set of all linear functionals of V be denoted by $\tilde{V}$ where we will also include the linear functional $y(x) = 0$ in $\tilde{V}$. If you look at the properties of a linear space on page 4-V, you can readily convince yourself that the set of all linear functionals $\tilde{V}$ forms a linear vector space. We call $\tilde{V}$ the "dual space" of V.

Now let us show that there is a one-to-one correspondence between V and $\tilde{V}$ for an n-dimensional space.

Theorem 5.1 *If $\{x_1, x_2, \ldots x_n\}$ is a basis for the n-dimensional vector space V and $\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is any set of n scalars, then there is one and only one linear functional y on V such that $y(x_i) = \alpha_i$, for $i = 1, 2 \ldots, n/$*

Proof Consider the vector $\sum_i \xi_i x_i$ where the ξ_i's are unique for each x. Then $\sum_i \xi_i y(x_i) = \sum_i \xi_i \alpha_i$. Can we now find another linear functional, say $y'(x)$ where y' is different from y but were $y'(x_i) = \alpha_i$?

If we can, we can write $y'(x) = \sum_i \xi_i' \alpha_i = \sum_i \xi_i' y'(x_i) = y'(\sum_i \xi_i' x_i)$ in which case $x = \sum_i \xi_i' x_i$. But since the x_i's form a basis and we had $x = \sum_i x i_i x_i$ above, we see that we must have $\xi_i' = \xi_i$ and so $y(x) = y'(x)$ ∎

Now we want to show that V and $\tilde{V}$ have the same dimensions. To do this, we consider the basis $\{x_1, x_2, \ldots, x_n\}$ in V and the set $y_1, y_2, \ldots, y_n$ in $\tilde{V}$ such that $y_i(x_j) = \delta_{ij}$.

Note that in doing this, we have picked out from all the linear functionals $\tilde{V}$ that set on n linear functionals where $y_i(x_j) = \delta_{ij}$ and we know that they are unique from our previous theorem. To show that these n linear functionals form a basis, we need to show that: 1) They are linearly independent, and 2) That they span $\tilde{V}$.

1. Can we have $\sum_i \alpha_i y_i(x) = 0$ for any x? Suppose $x = x_j$. Then we need $\alpha_j = 0$. Likewise, for $j = 1, 2, \ldots, n$. Thus all α_i's are zero and the y_i's are linearly independent.

2. Now we consider a linear functional y in $\tilde{V}$ such that $y(x_i) = \alpha_i$ for $i = 1, 2, \ldots, n$. Then for any x such that $x = \sum_i l, \xi_i x_i$, we have $y(x) = \sum_i \xi y(x_i) = \sum_i \xi \alpha_i$. Since $y_i(x_j) = \delta_{ij}$, we see that $y_i(x) = \xi_i$. Substituting this for ξ_i above, we have $y(x) \sum_i \alpha_i y_i(x)$. So all linear functionals y in $\tilde{V}$ can be written as a linear combinations of the y_i's. So the y_i's span $\tilde{V}$ and form a basis.

The space $\tilde{\tilde{V}}$ which is dual to $\tilde{V}$ is isomorphic to V. Shouldn't be too hard to prove.

5.8 The Dual Space 2

We consider the linear vector space V with the elements $|x\rangle, |y\rangle, |z\rangle \ldots$ which we will assume is a Hilbert space. This is the only kind of space we are interested in in quantum mechanics. "In" or "on" this vector space, V, we have a scalar product $\langle x|y\rangle$ defined for any two vectors $|x\rangle$ and $|y\rangle$ in V.

There is another linear vector space $\tilde{V}$, which is called the vector space "dual" to V or simply the "dual space". $\tilde{V}$ will have elements which consists of all the linear functionals (linear functionals) on V. We shall see that these is a one-to-one correspondence between the elements of V and the elements of $\tilde{V}$. We shall also see that the dimensions of V and $\tilde{V}$ are equal.

For simplicity in notation, I will use $x, y, z, \ldots$ for the vectors in V instead of $|x\rangle, |y\rangle, |z\rangle, \ldots$.

Definition A linear functional on a vector space V is a scalar-valued function $F(x)$ defined such that for every vector $x = \alpha_1 x_1 + \alpha_2 x_2$ in V, we have

$$F(x) = \alpha_1 F(x_1) + \alpha_2 F(x_2)$$

with α_1 and α_2 arbitrary complex numbers.

First we show that all linear combinations of linear functionals are themselves linear functionals. That is, if $F_1(x)$ and $F_2(x)$ are linear functionals, then so is $F(x) = \beta_1 F_1(x) + \beta_2 F_2(x)$. To show this, use $x = \alpha_1 x_1 + \alpha_2 x_2$. Then

$$F(x) = \beta_1 F_1(\alpha_1 x_1 + \alpha_2 x_2) + \beta_2 F_2(\alpha_1 x_1 + \alpha_2 x_2)$$

and as F_1 and F_2 are linear functionals,

$$\begin{aligned} F(x) &= \beta_1\{\alpha_1 F_1(x_1) + \alpha_2 F_1(x_2)\} + \beta_2\{\alpha_1 F_2(x_1) + \alpha_2 F_2(x_2)\} \\ &= \alpha_1\{\beta_1 F_1(x_1) + \beta_2 F_2(x_1)\} + \alpha_2\{\beta_1 F_1(x_2) + \beta_2 F_2(x_2)\} \\ &= \alpha_1 F(x_1) + \alpha_2 F(x_2) \end{aligned}$$

which completes the proof that $F(x)$ is a linear functional.

Looking at the properties of a linear vector space on pp 4-V, you can readily convince yourself that the set of all linear functionals on V ill form a linear vector space which we call the dual space, $\tilde{V}$.

Theorem 5.2 *For each linear functional F on V there is a unique vector, ψ_F, in V such that $F(x) = \langle \psi_F | x \rangle$.*

This theorem tells us that the value of the linear functional $F(x)$ is given by the scalar product $\langle \psi_F | x \rangle$ where ψ_F is unique.

To prove this, consider the orthonormal basis $\{x_1, x_2, x_3, \ldots\}$ in V. If there is such a vector ψ_F, then it is unique and we know what it is. If $F(x_i) = \langle \psi_F | x_i \rangle$ for every x_i, then we can write

$$\psi_F = \sum_{i=1}^{\infty} \langle x_i | \psi_F \rangle x_i = \sum_{i=1}^{\infty} F(x_i)^* x_i.$$

If $\sum_i F(x_i)^* x_i$ converges to a vector ψ_F, we have for any vector y that $y = \sum_{i=1}^{\infty} \langle x_i | y \rangle | x_i \rangle$ and so

$$F(y) = \sum_{i=1}^{\infty} \langle x_i | y \rangle F(x_i) = \langle \psi_F | y \rangle.$$

So ψ_F does the job if $\sum_{i=1}^{\infty} F(x_i)^* x_j$ converges to ψ_F. One can show that if F is continious and because it is linear, then this is indeed the case.

Theorem 5.3 *V and $\tilde{V}$ have the same dimensions.*

Proof To prove this we consider the orthonormal basis $\{x_1, x_2, x_3, \ldots, x_n\}$ in V and the set of linear functionals $\{y_1, y_2, \ldots, y_n\}$ in $\tilde{V}$ such that $y_i(x_j) =: \langle \psi_{y_i} | x_j \rangle = \delta_{ij}$ So we have picked a special set of linear functionals from $\tilde{V}$. Our previous theorem tells us that the vectors ψ_{y_i} in V unique. To show that this set of linear functionals in $\tilde{V}$ form a basis, we have to show: 1) They are linearly independent; and 2) They span $\tilde{V}$.

1. Can we have $\sum_i \alpha_i y_i(x) = 0$ for any x and some $\alpha_i \neq 0$? Consider $x = x_j$. Then $\sum_i \alpha_i y_i(x_j) = \alpha_j = 0$. Therefore $\alpha_j = 0$. Similarly we can prove that all α_i's $= 0$. For this set of y_i's This the y_i's are linearly independent.

2. Consider the above set $\{x_1, x_2, \ldots\}$ in V and the above set $\{y_1, y_2, \ldots\}$ in V. Also consider another linear functional y in $\tilde{V}$ such that $y(v_i) = \alpha_i$ for all x_i's. Then for any x such that $x = \sum_i \xi_i x_i$, we have

$$y(x) = \sum_i \xi_i y(x_i) = \sum_i \xi_i \alpha_i.$$

But since $y_i(x_j) = \delta_{ij}$, $y_i(x) = \xi_i$. Substituting this $y_i(x)$ for ξ_i, we have $y(x) = \sum_i \alpha_i y_i(x)$. This result tells us that we can express all linear functionals in $\tilde{V}$ as a linear combination of the y_i's. So the y_i's span $\tilde{V}$ and form a basis. As the number of x_i's equals the number of y_i's, V and $\tilde{V}$ have the same dimensions ∎

It shouldn't be too hard to show that the space $\tilde{\tilde{V}}$ dual to $\tilde{V}$ is isomorphic to the space V.

5.9 Dirac's Bra and Ket Notation 2

Dirac calls the vectors $|x\rangle, |y\rangle, |z\rangle, \ldots$ in V by the name of kets and the scalar product is $\langle x|y\rangle$ for two vectors $|x\rangle$ and $|y\rangle$.

Instead of writing $F(x) = \langle \psi_F | x \rangle$ as did previously, let us change notation and write $F_\psi(x) = \langle \psi | x \rangle$, i.e., to every linear functional F_ψ in $\tilde{V}$ we have a unique vector $|\psi\rangle$ in V. In fact, lets go one step further and change notation again and let $\langle \psi |$ be the vector in $\tilde{V}$ which denotes the vector F_ψ in $\tilde{V}$.

Going one step further, we combine the linear functional $\langle\psi|$ in $|tildeV$ with the vector $|x\rangle$ in V and use the notation that $\langle\psi|x\rangle$ denotes the linear functional $F_\psi(x)$. If we do this, we have no way of distinguishing the linear functional $\langle\psi|x\rangle$ in $\tilde{V}$ from the scalar product $\langle\psi|x\rangle$ in V. It really doesn't matter since they are equal.

But now we can use $|x\rangle, |y\rangle, \ldots$ for vectors in V and $\langle x|, \langle y|, \ldots$ for vectors in $\tilde{V}$.

Dirac calls the vectors $\langle x|, \langle y|, \ldots$ bras and $\langle x|y\rangle$ is called a "bra-ket".

I prefer the above approach to that of Messia's e.g., on pp 247-249 Vol. I, he has 6 places where he either hypothesizes, on assumes, or postilates.

The relationship between $|x\rangle$ and $\langle x|$ must be anti-linear. If $|x\rangle = \alpha_1|x_1\rangle + \alpha_2|x_2\rangle$ then

$$\begin{aligned} \langle z|x\rangle &= \alpha_1\langle z|x_1\rangle + \alpha_2\langle z|x_2\rangle \text{ while} \\ \langle x|z\rangle &= \langle z|x\rangle^* = \alpha_1^*\langle x_1|z\rangle + \alpha_2^*\langle x_2|z\rangle \text{ so} \\ \langle x| &= \alpha_1^*\langle x_1| + \alpha_2^*\langle x_2| \end{aligned}$$

giving the anti-linear relationship.

We now explore additional relationships involving the bras, kets, linear operators, scalars and scalar products.

Consider the linear operators A in V which changes $|y\rangle$ into $|x\rangle$, i.e., $|x\rangle = A|y\rangle$. We should like to find the corresponding operators, $\overline{A}$ in $\tilde{V}$ which changes $\langle y|$ into $\langle x|$, i.e., $\langle x| = \langle y|\overline{A}$. Instead of using A and $\overline{A}$, let us write these expressions as $|x\rangle = (A|y\rangle$ and $\langle x| = (\langle y|A)$ respectively where the parenthesis indicate in which space the A's are acting. Now the $\overline{A}$ and A need not be equal. $\overline{A}$ may not even be defined in V. If it was, $\overline{A}|y\rangle$ need not give $|x\rangle$. However, if A in V on $|y\rangle$ produces $|x\rangle$, then we want the corresponding conjugate operator A in $\tilde{V}$ on $\langle y|$ to produce $\langle x|$, i.e., if $|x\rangle = (A|y\rangle)$, we want $\langle x| = (\langle y|A)$. One can easily show that if one of the A's is a linear operator, then so is the other conjugate A.

Let us now show that $\langle z|(A|y\rangle)$ is a linear functional on $|y\rangle$. Using $|y\rangle = \alpha_1|y_1\rangle + \alpha_2|y_2\rangle$, we have $\langle z|(A|y\rangle)$ is a linear functional of $|y\rangle$ and we define a bra (linear functional) $\langle x|$ such that $\langle x|y\rangle =: \langle z|(A|y\rangle)$.

Is $\langle x|$ unique? If not, we can find two $\langle x|$'s, say $\langle x_1|$ and $\langle x_2|$, such that $\langle x_1|y\rangle = \langle z|(A|y\rangle)\rangle$ and $\langle x_2|y\rangle = \langle z|(A|y\rangle)\rangle$. Subtracting gives $\langle x_1|y\rangle - \langle x_2|y\rangle = (\langle x_1| - \langle x_2|)|y\rangle = 0$ for all $|y\rangle$. This can only happen if $\langle x_1| - \langle x_2| = 0$ or $\langle x_1| = \langle x_2|$ and so $\langle x|$ is unique.

For a given linear operator A, each different bra $\langle z|$ defines a different unique bra $\langle x|$. We will define

the linear operator A in $\tilde{V}$ as the one which when acting on $\langle z|$ will give $\langle x|$, i.e, $\langle x| = (\langle z|A)$. With this definition, we have

$$\langle x|y\rangle = (\langle z|A)|y\rangle) = \langle z|(A|y\rangle).$$

We see that we get the same scalar product independently of whether A in V operates on $|y\rangle$ or whether A in $\tilde{V}$ operates on $\langle z|$. In this case, one usually drops the parenthesis and writes $\langle z|A|y\rangle$ for both $(\langle z|A)|y\rangle$ and $\langle z|(A|y\rangle)$.

Dropping the parenthesis is okay for linear operators but will not be okay in general, e.g., for anti-linear operators, the parenthesis must be retained and it will make a difference whether A operates to the left of to the right.

$|x\rangle = A|y\rangle$ need not be unique. One could expect to find many A's and/or $|y\rangle$'s which give the same $|x\rangle$, e.g, rotation of all vectors of the same length into a given direction.

We now explore additional relationships involving the bras, kets, linear operators, and scalar products. $|x\rangle = A|y\rangle$ means that the linear operator A operates on $|y\rangle$ and changes it into $|x\rangle$. Since $|x\rangle$ and $|y\rangle$ have their conjugates in the dual space, we can ask for an operator associated with A in the dual space, say $\overline{A}$ which will change $\langle y|$ into $\langle y|$. That is, $\langle x| = \langle y|\overline{A}$ where $\overline{A}$ operates "to the left" on $\langle y|$. Instead of using this notation, let us write $|x\rangle = (A|y\rangle)$ and $\langle x| = (\langle y|A)$ where it is clear whether A operates in the vector space V or in the dual space $\tilde{V}$. One can easily show that if A is a linear operator in V, then the "associated" A in $\tilde{V}$ is also a linear operator.

Consider the linear operator A and any bra $\langle z|$. Let is show that there is a unique bra $\langle x| =: (\langle z|A)$ such that for all kets $|y\rangle$, we have $\langle x|y\rangle = \langle z|(A|y\rangle)$.

To show this, suppose that there are two bras $\langle x_1|$ and $\langle x_2|$. If so, we have

$$\langle x_1|y\rangle = \langle z|(A|y\rangle) \text{ and } \langle x_2|y\rangle = \langle z|(A|y\rangle).$$

Subtracting, we have

$$\langle x_1|y\rangle - \langle x_2|y\rangle = (\langle x_1| - \langle x_2|)|y\rangle = 0 \text{ for all } |y\rangle.$$

But this can be true $\iff \langle x_1| - \langle x_2| = 0$ or $\langle x_1| = \langle x_2|$. So by defining this unique bra $\langle x|$ as $(\langle z|A)$, we have shown that

$$\langle x|y\rangle = (\langle z|A)|y\rangle = (\langle z|(A|y\rangle),$$

which tells us that our scalar produce is independent of whether A operates to the right on $|y\rangle$ or to the left on $\langle z|$. In this case, one usually "drops" the parenthesis and writes both $(\langle z|A)|y\rangle$ and $\langle z|(A|y\rangle)$ as $\langle z|A|y\rangle$.

Now this is fine for linear operators. For anti-linear operators, (e.g., the time reversal operator) where $A(\alpha|x\rangle = \alpha^*(A]ketx)$, one is not able to drop the parenthesis as it will make a difference whenther A operates to the right or to the left. RRAY Note that $|x\rangle = A|y\rangle$ need not be unique as one could expect to find many $|y\rangle$'s and/or many A's which give the same $|x\rangle$. For example, rotation of all vectors of same length into a given direction.

Another linear operator is $A = |u\rangle\langle u|$. In this case,

$$\begin{aligned} A|x\rangle &= |u\rangle\langle u|x\rangle \text{ and} \\ A|y\rangle &= |u\rangle\langle u|y\rangle \end{aligned}$$

If $\langle u|x\rangle = \langle u|y\rangle$, the "projection" of $|x\rangle$ and $|y\rangle$ on $|u\rangle$ is the same. We see that $A|x\rangle = A|y\rangle$ and so A acting on $|x\rangle$ and $|y\rangle$ produces the same vector.

We note that $A|x|$ going to $|z\rangle$, i.e, $A|x\rangle = |z\rangle$, may be unique while going from $|z\rangle$ back to $A|x\rangle$ may not be unique. Just above, we have a good example of this when $A|x\rangle = A|y\rangle = |u\rangle\langle u|x\rangle = |u\rangle\langle u|y\rangle$ with $\langle u|x\rangle = \langle u|[\rangle y\rangle$. However, if $|x\rangle = A|y\rangle$ and $B|x\rangle = |y\rangle$ for all x and y, we ss that $AB = 1$ and so A and B are "inverses" of each other. Usually we write A^{-1} for B or B^{-1} for A. One can easily show that $(AB)^{-1} = B^{-1}A^{-1}$ – reverse order as we would expect.

We can also approach inverses in a little different manner. Let A have the two properties

1. If $[|x_1\rangle \neq |x_2\rangle$, then $A|x_1\rangle \neq A|x_2\rangle$ for all $|x_1\rangle$ and $|x_2\rangle$.

2. To every vector $|y\rangle$ there corresponds at least one vector $|x\rangle$ such that $A|x\rangle = |y\rangle$.

If A has these two properties, A is said to be "invertible" and one can find the inverse A^{-1}. In finite dimensional vector spaces, 1. implies 2., but this is not the case for infinite dimensional vector spaces.

This we see that while $|u\rangle\langle u|$ is a linear operator, it does not possess an inverse. If A does not possess and inverse, then A has zero eigenvalues. If $A|x\rangle = A|y\rangle$, then $A(|x\rangle - |y\rangle) = 0$.

If we know how certain operators go with the kets, we can now state how they will go with the bras. For example

$$\begin{array}{rlcrl} \text{If } (cA)|x\rangle &= C(A|x\rangle & \text{then} & \langle x|(AC) &= (\langle x|A)C \\ (\alpha A)|x\rangle &= \alpha(A|x\rangle & & \langle x|(A\alpha) &= \alpha(\langle x|A) \\ (A+B)|x\rangle &= A|x\rangle + B|x\rangle & & \langle x|(A+B) &= \langle x|A + \langle x|B \\ A(B|x\rangle) &= P|x\rangle & & (\langle x|(A)B &= \langle x|P \end{array}$$

5.10 Hermitian Operators

Consider the ket $|\alpha\rangle$ which is conjugate to the bra $\langle\alpha| =: \langle\beta|A$. Now for every bra $\langle\alpha|$ and every bra $\langle\beta|$, we can find the corresponding kets $|\alpha\rangle$ and $|\beta\rangle$ respectively. With this in mind, we define the linear operator $A^\dagger$ such that if $\langle\beta|A = \langle\alpha|$, then $A^\dagger|\beta\rangle = |\alpha\rangle$. We call $A^\dagger$ the "Hermitian conjugate" or "Hermitian adjoint" of A.

With this definition of $A^\dagger$, we can show several useful things.

1. First, let us show that this $A^\dagger$ is consistent with the $A^\dagger$ we introduced on page 10-I in connection with scalar products in the coordinate representation which was $(\phi, A\psi) = (A^\dagger\phi, \psi) = (\psi, A^\dagger\phi)^*$.

Here, let $|x\rangle = A|y\rangle$ and then $\langle x| = \langle y|A^\dagger$. Using these in $\langle z|x\rangle = \langle x|z\rangle^*$, we have

$$\langle z|A|y\rangle = \langle y|A^\dagger|z\rangle^*$$

and it is consistent with our previous definitions.

2. $(A^\dagger)^\dagger = A$.

Proof Now $\langle x|A|y\rangle = \langle y|A^\dagger|x\rangle^*$ where A and $A^\dagger$ can operate either to the right or to the left. As $|x\rangle$ is the conjugate of $\langle x|$, then $A|y\rangle$ is the conjugate of $\langle y|A^\dagger$. But from above, we know that the conjugate of $\langle y|A^\dagger$ is $(A^\dagger)^\dagger|y\rangle$. This we see that $(A^\dagger)^\dagger = A$.

3. $(\alpha A)^\dagger = \alpha^* A^\dagger$

Proof Using $(\alpha A)^\dagger$ for A in case 2 above, we have

$$\begin{aligned}\langle x|(\alpha A)^\dagger|y\rangle &= \langle y|(\alpha A|x\rangle^* = \alpha^*\langle y|A|x\rangle^* = \alpha^*\langle x|A^\dagger|y\rangle \\ &= \langle x|\alpha^* A^\dagger|y\rangle \quad \blacksquare\end{aligned}$$

$$\left.\begin{aligned}4.(A+B)^\dagger &= A^\dagger + B^\dagger \\ 5.(AB)^\dagger &= B^\dagger A^\dagger\end{aligned}\right\} \text{ Proofs easy. Relations reasonable}$$

6. $(|u\rangle\langle v|)^\dagger = |v\rangle\langle u|$

Proof

$$\begin{aligned}\langle x|(|u\rangle\langle v|)^\dagger|y\rangle &= (\langle y|(|u\rangle\langle v|)|x\rangle)^* = (\langle y|u\rangle\langle v|x\rangle)^* \\ &= \langle x|v\rangle\langle u|y\rangle = \langle x|(|v\rangle\langle u||y\rangle.\end{aligned}$$

The above relationships allow us to give a general rule for taking the Hermitian conjugate of an expression involving numbers, operators, bras and kets. One replaces α by α^*, A by $A^\dagger$, $|x\rangle$ by $\langle x|$, $\langle y|$ by $|y\rangle$ and writes all the factors in reverse order.

For example:

$$\left(A|x\rangle\langle y|\alpha B^\dagger C|z\rangle\right)^\dagger = \langle z|C^\dagger B\alpha^*|y\rangle\langle x|A^\dagger.$$

If $A^\dagger = A$, A is called a "Hermitian operator". If $A^\dagger = -A$, A is called "anti-Hermitian"or "skew Hermitian".

If $A^\dagger = A^{-1}$ so that $AA^\dagger = 1$, A is called an "Unitary operator".

5.11 Eigenvalues and Eigenspectra

If we have $A|u\rangle = a|u\rangle$ where a is a complex number, we say that $|u\rangle$ is an eigenket of the operator A and a is the eigenvalue. Similarly for $b\langle u| = \langle u|B$, we have $\langle u|$ as the eigenbra of the operator B and b is the eigenvalue.

Several indpendent kets can have the same eigenvalue a. These kets form a subspace and the dimensions of this subspace is called the order of the degeneracy. Similar statements can be made about b and the eigenbras.

In general, the eigenvalue problem

$$A|u\rangle = a|u\rangle$$

and the eigenvalue problem

$$\langle u|A = a'\langle u|$$

are quite different problems since the later one is equivalent to solving

$$A^\dagger|u\rangle = a'^*|u\rangle a''|u\rangle$$

and A and $A^\dagger$ are different operators.

However, for a Hermitian operator $A^\dagger = A$, we have

1. The eigenspectra are identical

2. All eigenvalues are real.

3. Each bra conjugate to an eigenket is an eigenbra with the same eigenvalue.

4. The eigenkets belonging to different eigenvalues are orthogonal.

In general, the eigenkets of a Hermitian operator can be split up into those kets with discrete eigenvalues

$$A|n\rangle = a_n|n\rangle$$

and those kets with continuous eigenvalues

$$A|\nu\rangle = a(\nu)|\nu\rangle.$$

If r of the $|n\rangle$'s were degenerate (same a_n), we can use the "Schmidt process" to concoct r orthonormal functions. If we labeled these functions by $|n, r\rangle$, we clearly have

$$\langle n, r||n', r'\rangle = \delta_{nn'}\delta_{rr'}.$$

Similarly in the continuous eigenvalue case, the degeneracy could be labeled by some discrete index (or indices), r, taking on a finite or infinite number of values, by an index , α, (or indices) varying continuously, or by a combination of these. In this case, we have

$$\langle \nu, \alpha, r|\nu', \alpha', r'\rangle = \delta(\nu - \nu')\delta(\alpha - \alpha')\delta_{rr'}$$

and clearly

$$\langle n, r|\nu, \alpha, s\rangle = 0.$$

If the eigenkets $|n, r\rangle$ and $|\nu, \alpha, s\rangle$ span our entire vector space, we say that they form a complete set and the Hermitian operator A is called an "observable".

In general, it is difficult to prove that we have a complete set and so it is often difficult to prove that a given operator is an observable, Usually physicists are guided by classical mechanics and "good judgement".

5.12 Projection Operators

Let us split up our Hilbert Space, HS, into two subspaces $\mathcal{L}_s$ and $\mathcal{L}_a$ where $\mathcal{L}_s + \mathcal{L}_a = HS$. For example, we know that the bound states of the hydrogen atom can be described by the eigenkets $|n, J, m, \pi\rangle$. We could specify $\mathcal{L}_s$ as the subspace covered by all those eigenkets with a given value of J. $\mathcal{L}_a$ would then consist of all the rest.

In general, any ket $|u\rangle$ can be written as a sum of two kets – one in $\mathcal{L}_s$ and one in $\mathcal{L}_a$, i.e., $|u\rangle = |u_s\rangle + |u_a\rangle$, where clearly this "breakup" is unique.

Consider the linear operator P_s, called a "projection" operator which will project out the $\mathcal{L}_f$ part, i.e., $P_s|u\rangle = |u_s\rangle$. Clearly $P_s|u_s\rangle = |u_s\rangle$ and $P_s|u_a\rangle = 0$.

We first show that P_s is a Hermitian operator by using $\langle u_s|u_a\rangle = 0$ for any $|u\rangle$ and $|v\rangle$.

Now $\langle u|P_s|v\rangle = \langle u|v_s\rangle = \langle u_s|v_s\rangle = \langle u_s|v\rangle$.

Thus $\langle u|P_s = \langle u_s|$ or $P_s^\dagger|u\rangle = |u_2\rangle = P_s|u\rangle$.

Since $P_s^2|u\rangle = P_s|u\rangle$, we have $P_s^2 = P_s$ and $P_s(P_s - 1)|u\rangle = 0$. This latter is an eigenvalue equation which tells us that P_s had the eigenvalue of 0 or 1. If it is 1, $|u\rangle = |u_s\rangle$ and if it is 0, $|u\rangle = |u_a\rangle$.

Since $(1 - P_s)|u\rangle = |u_a\rangle$, we see that $P_a = 1 - P_s$ or $P_a + P_s = 1$.

As a simple example, suppose that $\mathcal{L}_s$ consisted of a single ket, $|s\rangle$ – one dimensional – and let $|u_s\rangle$ be the projection of $|u\rangle$ on $\mathcal{L}_s$.

We have

$$\begin{aligned} P_s|s\rangle &= |s\rangle \\ P_s|u\rangle &= |u_s\rangle \\ (1 - P_s)|u\rangle &= |u_a\rangle \\ \langle u_s|s\rangle &= 0. \end{aligned}$$

Now $\langle s|u_s\rangle =$ same complex $\# =: \alpha$, or $|u_s\rangle = \alpha|s\rangle$ if $\langle s|s\rangle = 1$. Multiplying $\langle s|u\rangle$ by $|s\rangle$, we have

$$|s\rangle\langle s|u\rangle = |s\rangle\langle s|u_s\rangle = \alpha|s\rangle = |u_s\rangle = P_s|u\rangle.$$

As this is true for all $|u\rangle$, we see that

$$P_s = |s\rangle\langle s|.$$

We can generalize this in the same manner. Consider the subset, $\mathcal{L}_s$, of the linearly independent kets, $|n\rangle$, with $n = 1, 2, 3, \ldots, N$ where $\langle n|m\rangle = \delta_{nm}$. In this case, the projection operator becomes

$$P_s = \sum_{n=1}^{N} |n\rangle\langle n|.$$

For a subset of kets with continuous indicese with $\langle \nu|\nu'\rangle = \delta(\nu - \nu')$, the projection operator becomes (suppressing the "s" subscript)

$$P = \int |\nu\rangle \, d\nu \, \langle \nu|.$$

For a subspace with degenerate and discrete eigenkets labeled by $|n, r\rangle$, we have

$$P = \sum_{n,r} |n, r\rangle\langle n, r|.$$

If our subspace was the whole of Hilbert space, P becomes unity and we have

$$\sum_{n,r} |n, r\rangle\langle n, r| = 1.$$

This relation is called a "closure relation" or a "reduction of unity". The closer relation for continuous indices with degeneracies labeled by r is

$$1 = \sum_{r} \int |\nu, r\rangle \, d\nu \, \langle \nu, r|.$$

For a discrete and continuous case, our closure relation is

$$1 = \sum_{n,r} |n, r\rangle\langle n, r| + \sum_{r} \int |\nu, r\rangle \, d\nu \, \langle \nu, r|.$$

These projection operators are closure relations are very important and useful. In particular, we will make great use of the closure relations when we discuss the matrix representations as we go from the abstract Hilbert space into a given representation. The coordinate representation is the Schrödinger picture and the one most used in quantum mechanics.

We now turn to a quick review of matrices and then discuss matrix representations and Unitary transformations from one representation to another.

5.13 Matrices and the Matrix Representation

Finite Matrices

It is often more convenient to express the solutions of quantum mechanical problems by using matrix notation instead of vectors in our abstract Hilbert space. In fact, we shall see that the Schrödinger picture is really nothing more than a matrix representation called the coordinate representation.

For completeness, we write down some of the more important properties of finite matrices – properties which you probably know.

A M-by-N matrix is a rectangular array of M rows and N columns. For example,

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{i1} & A_{i2} & \dots & A_{in} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \tag{1}$$

If $M = N$, we have a "square" matrix.
If $N = 1$ and $M \neq 1$, we have a "column" matrix.
If $N \neq 1$ and $M = 1$, we have a "row" matrix.
If $M = N = 1$, we have a 1-by-1 matrix called a "scalar".
We shall see that the kets can be associated with a column matrix, the bras with a row matrix, and operators as square matrices.

In the properties listed below let A_{nm} be the n-th row, m-th column element.

1. If $A_{nm} = B_{nm}$ for all n and m, then $A = B$.
2. $(\alpha A)_{nm} = \alpha A_{nm}$.
3. $S = A + B = B + A$ where $S_{nm} = A_{nm} + B_{nm}$.
4. $P = AB$ where $\sum_k A_{nk} B_{km}$ and clearly the number of columns of A must equal the number of rows of B for this to make sense.
5. There exists a null matrix O, where $OA = AO = O$ and $O_{nm} = 0$.
6. There exists a unit matrix I, where $I_{nm} = \delta_{nm}$ and $IA = AI = A$.
7. $\tilde{A}$ = transpose of A such that $(\tilde{A})_{nm} = A_{mn}$.
8. A^* = complex conjugate of A where $(A^*)_{nm} = (A_{nm})^*$.
9. $A^\dagger$ = Hermitican conjugate of A where $(A^\dagger)_{nm} = (A_{mn})^* = A^*_{mn}$ (note reverse order of indices).

Some special properties for square matrices are:

10. Trace of $A = \mathrm{Tr}A = \sum_{n=1}^{N} A_{nn}$ and the trace of a product of matrices is invariant to cyclic permutations., i.e., $\mathrm{Tr}(ABCD) = \mathrm{Tr}(BCDA)$.
11. The determinant of a matrix $A = \det A$ is the determinant of the square array. Also, $\det(ABC) = (\det A)(\det B)(\det C)$. A determinant is "singular" if $\det A = 0$ and "nonsingular" if $\det A \neq 0$.
12. If $AB = I$ or $BA = I$, B is called the inverse of A and written as A^{-1}. For A^{-1} to exist, A must be nonsingular and $(A^{-1})_{nm} =$ Cofactor of $A_{mn}/(\det A)$.
13. A is called an "orthogonal matrix" if $\tilde{A} = A^{-1}$. Then $\tilde{A}A = A\tilde{A} = 1$.
14. A is called a "unitary matrix" if $A^\dagger = A^{-1}$. Then $A^\dagger A = AA^\dagger = 1$.

From these definitions follow other properties – some of which are:

$$\begin{aligned}
(AB)^* &= A^*B^* \\
\tilde{AB} &= \tilde{B}\tilde{A} \\
(AB)^\dagger &= B^\dagger A^\dagger \\
(AB)^{-1} &= B^{-1}A^{-1} \\
(\tilde{A})^{-1} &= (\tilde{A^{-1}}) \\
(A^*)^{-1} &= (A^{-1})^* \\
(A^{-1})^\dagger &= (A^\dagger)^{-1}.
\end{aligned}$$

We now prove two statements.
1. If A is singular, one can find a $|u\rangle \neq 0$ such that $A|u\rangle = 0$ and conversely, if $A|u\rangle = 0$ for $|u\rangle =\neq 0$, then A is singular.

Proof We prove the converse first. Let $|u\rangle$ be a column matrix with elements $u_1, u_2, \ldots, u_N$. Let A be an N-by-N matrix and consider each column in the matrix as a vector. We can then write $A = (\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_N)$ as a row matrix where $(\underline{a}_i)_n =: A_{ni}$. Then $A|u\rangle = \sum_i \underline{a}_i u_i$ and we see that $A|u\rangle$ is a linear combination of the N $\underline{a}_i$ vectors. If $A|u\rangle = 0$ and not all $u_i = 0$, then the N a_i vectors are linearly dependent and so at least one of them is a linear combination of the others. This is enough to make $\det A = 0$ which proves the converse.

The converse to the converse is that if A is singular, the $\underline{a}_i$'s are linearly dependent and so one can find a $|u\rangle \neq 0$ such that $A|u\rangle = 0$. ∎

2. If A is non-singular and $A|u\rangle = 0$, then $|u\rangle = 0$. Conversely, if $A|u\rangle = 0$ implies $|u\rangle = 0$, then A is non-singular.

Proof We prove the converse first. If $A|u\rangle = 0$ implies that $|u\rangle = 0$, then the $\underline{a}_i$'s above a linearly independent and can be used as a basis. Now we consider the special basis $\underline{e}_1.\underline{e}_2, \underline{e}_3, \ldots, \underline{e}_N$ where $(\underline{e}_i)_j = \delta_{ij}$. Next we expand the $\underline{e}_i$'s in terms of the $\underline{a}_j$'s, i.e. $\underline{e}_i = \sum_j B_{ji}\underline{a}_j$ or in component form $\delta_{ik} = \sum j, k\, A_{kj}B_{ji}$. The above tells us that B is the inverse of A and so A must be non-singular.

Conversely, if A is non-singular, then B exists and in order to have $A|u\rangle = 0$, we need $|u\rangle = 0$. ∎

These two proofs tell us that if we have two square matrices A and B and a vector $|u\rangle \neq 0$, then in order to have $A|u\rangle = \alpha B|u\rangle$, it is necessary and sufficient that we have $\det(A - \alpha B) = 0$. All α's satisfying the above condition are solutions.

If $B = I$, then we want to solve the eigenvalue equation $A|u\rangle = \alpha|u\rangle$ and we see that we need $\det(A - \alpha I) = 0$. The equation $\det(A - \alpha I) = 0$ is called the "secular equation" and its roots – the allowed values of α – will be the eigenvalues of A. One usually runs into the secular equation when one wants to find the eigenvalues of $H|\psi_n\rangle = E_n|\psi_n\rangle$.

5.14 Infinite Matrices

Practically all of the properties of finite matrices carry over for infinite matrices. Infinite matrices are important for us as we will run into many of them in quantum mechanics. Our rows and columns will be labeled by a combination of finite or denumerably infinite number of discrete indices and an number of continuous indices which can take on all values in a given interval.

A matrix is said to be a square matrix if the rows and columns are labeled the same. If they are labeled differently, the matrix is considered to be a rectangular matrix.

In order to form the matrix product AB, the columns of A must be labeled the same as the rows of B. When we have continuous indices, the sums must be replaced by an integral. As an example, let us consider the square matrices A and B, which depend on the continuous index q where $q_1 \leq q \leq q_2$. The product of A and B, P is given by

$$P(q, q') = \int_{q_1}^{q_2} A(q, q'')B(q'', q')\, dq''$$

where we assume that the integral is finite.

For continuous indices, a diagonal matrix is defined as

$$D(q, q') = f(q)\delta(q - q').$$

This definition of a diagonal matrix preserves our "diagonal matrix properties" for finite matrices. For example, diagonal matrices commute with each other and multiplying a vector by this diagonal matrix just multiplies the "vector components" by the corresponding diagonal matrix element, i.e.

$$\int D(q, q')|x(q'\rangle\, dq' = f(q)|x(q)\rangle.$$

For infinite matrices, the concept of a a determinant is no longer valid.

For infinite matrices, we need both $AB = I$ and $BA = I$ for B to be the inverse of A. In additional A need not be a square matrix to possess an inverse. For example, A could have rows labeled by discrete indices and columns labeled by continuous indices. Then A^{-1} would have rows labeled by continuous indices and columns labeled by discrete indices. The two unit matrices in

$$AA^{-1} = I \text{ and } A^{-1}A = I$$

would be labeled differently in this case.

5.15 The Connection Between Matrices, Operators, and Vectors

Now we want to show how the operators and vectors which we have in quantum mechanics are connected to matrices. For simplicity, we will only consider discrete indices – but everything we do can easily be carried over for the case of continuous indices.

Let us assume that we know the eigenkets, $|n\rangle$ of some observable Q. From our definitions of an observable on page 48, we know that these eigenkets form a complete set and span our Hilbert space. For simplicity, we will assume that these eigenkets form an orthonormal basis. These eigenkets will be called the "basis vectors in the Q-representation."

In this case, our closure relation is

$$P_Q = \sum_n |n\rangle\langle n| = I,$$

and any vector $|u\rangle$ can be written as

$$|u\rangle = \sum_n |n\rangle\langle n|u\rangle.$$

We will regard the numbers, $|u_n\rangle\langle n|u\rangle$ as the elements of a column matrix. Then we would write

$$|u\rangle = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \end{pmatrix}$$

It is important to note that given the eigekets, $|n\rangle$, of the Q-representation and the u_i's then the vector $|u\rangle$ is uniquely defined.

Until now, all we have been doing is to talk about vectors in an abstract vector space. Now we have written down the components of these vectors. –And the components depend on the representation.

Similarly for the $\langle v|$, we have

$$\langle v| = \sum_n \langle v|n\rangle\langle n|$$

and we regard $v_n^* = \langle v|n\rangle = \langle n|v\rangle^*$ as the elements of a row matrix $\langle v| = (v_1^*, v_2^*, \ldots)$.

Since $u_n = \langle n|u\rangle$ and $u_n^* = \langle u|n\rangle$, we see that $|u\rangle^\dagger = \langle u|$ or

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \end{pmatrix}^{]} \; dagger = (u_1^*, u_2^*, \ldots).$$

We can "play the same game" for the linear operator A. THen

$$A = \sum_{n,m} |n\rangle\langle n|A|m\rangle\langle m|$$

and we call $A_{nm} = \langle n|A|m\rangle$ the matrix representation of the operator A in the Q-representation.

Likewise, for $A^\dagger$, we have

$$A^\dagger = \sum_{n,m} |n\rangle\langle n|A^\dagger|m\rangle\langle m| \text{ and } A^\dagger_{nm} = \langle n|A^\dagger|m\rangle.$$

Since $\langle n|A^\dagger|m\rangle = \langle m|A|n\rangle^*$, we see that

$$A^\dagger_{nm} = (A_{mn})^* = A^*_{mn}.$$

Note that this latter relationship is the same definition of the Hermitian conjugation in matrices. In fact, we are trying to show that there is a one-to-one correspondence betweetn all matrix operations and operations in our linear space. We have see in the matrix representations that the ket vectors are column matrices, operators are square matrices, and bra vectors are row matrices and are the Hermitian conjugate of the column matrix of the corresponding ket vector.

Some furhter examples of this correspondence are:

1. $\langle u|v\rangle = \sum_n \langle u|n\rangle\langle n|v\rangle = \sum_n u^*_n v_n$ which is just

$$(u^*_1, u^*_2, \ldots) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}.$$

2. A on $|u\rangle$

$$|x\rangle = A|u\rangle = \sum_{n,m} |n\rangle|xAynAm\langle m|u\rangle = \sum_n \left(\sum_m A_{nm} u_m \right) |n\rangle$$

and so $x_n = \sum_m A_{nm} u_m$.

3. A on $\langle v|$

$$\langle x| = \langle v|A = \sum_{n,m} \langle v|m\rangle\langle m|A|n\rangle\langle n| = \sum_m \langle n| \left(\sum_m v^*_m A_{mn} \right)$$

and so $x^*_n = \langle x|n\rangle = \sum_m v^*_m A_{mn}$.

4. $|x\rangle = AB|u\rangle = \sum_{n,l,m} |n\rangle\langle n|A|l\rangle\langle l|B|m\rangle\langle m|u\rangle$ and so $x_n = \sum_{l,m} A_{nl} B_{lm} u_m$.

One can continue on in this manner. However, we can already see that all the operations in our linear vector space carry directly over into matrix operations once we pick a representation.

But what is more important is that our formulation of quantum mechanics in terms of linear operators acting in our abstract Hilbert space is completely independent of a representation. That is, we could have picked another observable, say P, and use the eigenkets of P as a basis with which to express our ket vectors, bras, and linear operatos.

Even though we have used discrete indices in the above treatment, there is no problem in carrying everything over to continuous indices if our observable had eigenkets which were labeled in a continuous manner. The Schrödinger picture is the case where we use the coordinate representation. As we shall see later, the eigenkets in one dimension are $|q\rangle$ where $-\infty < q < \infty$. The for the state vector ψ, we have

$$\psi = \int |q\rangle\langle q|\psi\rangle \, dq \text{ as } 1 = \int |q\rangle \, dq \, \langle q|$$

and we define $\psi(q) \equiv \langle q|\psi\rangle$ which is a function of q and is the wave function we use in the Schrödinger picture. So the Schrödinger picture is just one of the <u>many</u> representations we could use to describe and/or solve quantum mechanical problems.

5.16 <u>Unitary Transformations</u>

A transformation is a very useful and a fundamental operation which can be thought of in one of two ways.

1. It can be considered as an operator which acts on a vector and changes this vector into a new vector. A simple example is an operator which rotates all vectors through a given angle around some given axis.

2. It can be considered as an operator which transforms from one basis to another basis, or one which transforms a vector expressed in terms of one basis to the same vector in terms of another basis. A simple example is the operator which transforms a vector expressed in a cartesian coordinate basis to one expressed in a spherical coordinate basis.

Now we usually are dealing with an equation (or equations) which contain operators, bras, kets, etc. and we can ask what will this equation look like after the transformation takes place.

In Case 1 above, we can always work with a given basis and given this basis, the transformation operator will have a matrix representation. However, in Case 2 above, the transformation operator will not have a matrix representation in a given representation because this transformation "con-

nects" two different bases. – Or "straddles" two different bases. But we will see that these two cases will behave formally in the same manner.

1. Matrix Transformations

For simplicity, let us consider matrix transformation with a matrix T which possesses an inverse and let u_r be a row matrix, u_e be a column matrix and A be a square matrix.

Consider the transformation defined by

$$\begin{array}{rlcrl} A' &= TAT^{-1} & \text{which implies that} & A &= T^{-1}A'T \\ U'_C &= TU_C & " \quad " \quad " & U_C &= T^{-1}U'_C \\ U'_r &= U_rT^{-1} & " \quad " \quad " & U_r &= U'_rT \end{array}$$

This transformation will preserve the form of the following algebraic matrix equations as well as the trace and the determinant,

$$\begin{array}{rlcrl} \text{If } A &= \alpha B + \beta CD & \text{then} & A' &= \alpha B' + \beta C'D' \\ U_r &= \alpha BV_r s_c A t_r & & U'_r &= \alpha B'V'_r s'_c A' t'_r \\ \text{and } \mathrm{Tr}A' &= \mathrm{Tr}A & \text{and} & \det A' &= \det A \end{array}$$

However, when our relations have Hermitian conjugates, we will have to put a further restriction on T if we want to preserve the form of the algebraic equations. Suppose we have $A = B^\dagger C$. It is reasonable to require that we have $A' = B'^\dagger C'$. If this is the case, it requires that $B'^\dagger = TB^\dagger T^{-1}$! But our original $B' = TBT^{-1}$ also requires that $(B^\dagger)' = TB^\dagger T^{-1}$ which tells us that $(B')^\dagger = (B^\dagger)'$. Since $(T^\dagger))^{-1} = (T^{-1})^\dagger$ (page 52), we can write

$$B' = ((B')^\dagger)^\dagger = (TB^\dagger T^{-1})^\dagger = (T^{-1})^\dagger BT^\dagger = (T^\dagger)^{-1}BT^\dagger.$$

Comparing this with $B' = TBT^{-1}$, we have $TBT^{-1} = (T^\dagger)^{-1}BT^\dagger$ or $T^\dagger TB = BT^\dagger T$ or $\left[{}^\dagger T, B\right] = 0$. For an operator, $T^\dagger T$ in this case, to commute with all B, it is necessary (stated without proof), that $T^\dagger T = \alpha I$, i.e, $T^\dagger T$ is a multiple of the unit matrix.

By demanding that $(U'_C)^\dagger = (U^\dagger_C)'$, we can find α. Now $U'_C = TU_C$ and so $(U'_C)^\dagger - U^\dagger_C T^\dagger$. Now $U^\dagger_C$ behaves like a row matrix which requires that $(U^\dagger_C)' = U^\dagger_C T^{-1}$. Comparing these last two expressions tell us that we have $T^\dagger = T^{-1}$.

So $T^\dagger T = TT^\dagger = I$ and $\alpha = 1$. We see that in order to have all algebraic equations preserved, we need to have an Unitary Tranformation.

Then

$$\left.\begin{array}{l} |u'\rangle = R|u\rangle \\ \langle v'| = \langle v|T^\dagger \\ A' = TAT^\dagger \end{array}\right\} \text{ where } TT^\dagger = T^\dagger T = I.$$

1. Matrix Transformations a little differently/.

As a side point, we can also approach our matrix transformations in a little different manner and arrive at the same conclusion. Consider the matrix transformation for column matrices

$$|u'\rangle = T|u\rangle.$$

When we make a measurement in the physical world we "measure" either $\langle u|v\rangle$ and/or $\langle u|A|v\rangle$. This is is reasonable to require that

$$\langle u'|v'\rangle = \langle u|v\rangle \text{ and } \langle u'|A'|v'\rangle = \langle u|A|v\rangle$$

since our measurements should not depend on the transformation T.

Now $|u'\rangle = T|u\rangle$ implies that $\langle u'| = \langle u|T^\dagger$. Then $\langle u'|v'\rangle = \langle u|v\rangle$ can be written as

$$\langle u'|v'\rangle = \langle u|T^\dagger T|v\rangle$$

and we see that we need $T^\dagger T = I$ and so T is required to be an Unitary Transformation.

Next, we can ask for the relation between A and A' when $\langle u'|A'|v'\rangle = \langle u|A|v\rangle$. Then we have $\langle u'|A'|v'\rangle = \langle u|T^\dagger AT|v\rangle$ and we see that $T^\dagger A'T = A$ or that $A' = TAT^\dagger$ which is the same as we had before. Furthermore we see that the form of all algebraic expressions as well as traces and determinants will be preserved.

2. Change of Representations

Now let us look at the 2nd case where we consider a transformation between two bases.

Let us consider two observables Q and P. Let Q have the eigenkets $|n\rangle$ and P have the eigenkets $|\overline{m}\rangle$ where we will use discrete indices for simplicity.

We can expand each basis in terms of the other. That is

$$|n\rangle = \sum_{\overline{m}} |\overline{m}\rangle\langle\overline{m}|n\rangle = \sum_{\overline{m}} |\overline{m}\rangle T_{\overline{m}n}$$

which defines T and

$$|\overline{m}\rangle = \sum_{n} |n\rangle\langle n|\overline{m}\rangle = \sum_{n} |n\rangle \overline{m} S_{n\overline{m}}$$

which defines S.

Since $\langle\overline{m}|n\rangle = \langle n|\overline{m}\rangle^*$, we have

$$T_{\overline{m}n} = S^*_{n\overline{m}} \text{ or that } T = S^\dagger \text{ or } S = T^\dagger.$$

Next we observe that

$$\langle \overline{m}|\overline{m}'\rangle = \sum_n \langle \overline{m}|n\rangle\langle n|\overline{m}'\rangle = \delta_{\overline{m},\overline{m}'}$$

which tells us that

$$\sum_n T_{\overline{m}n} S_{n\overline{m}'} = \delta_{overlinem,\overline{m}'}$$

or that $TS = I$ or $TT^\dagger = S^\dagger S = I$. Thus we see that our transformation from one basis to another basis is also an Unitary Transformation. Remember that our S's and/or T's are defined with respect to two given representations.

Now we look at an arbitrary vector $|u\rangle$.

Let $|u\rangle_Q$ refer to the vector $|u\rangle$ expressed in the Q representation.

Let $|u\rangle_P$ refer to the vector $|u\rangle$ expressed in the P representation.

The nth component of $|u\rangle_Q$ can be written as

$$\left(|u\rangle_Q\right)_n = \langle n|u\rangle = \sum_{\overline{m}} \langle n|\overline{m}\rangle\langle \overline{m}|u\rangle = \sum_{\overline{m}} T^\dagger_{n\overline{m}} \left(|u\rangle_P\right)_{\overline{m}}$$

or symbollically

$$|u\rangle_Q = T^\dagger |u\rangle_P = S|u\rangle_P.$$

Next we consider the linear operator A.

$$\begin{aligned}\left(A_Q\right)_{nm} &= \langle n|A|m\rangle = \sum_{\overline{k},\overline{l}} \langle n|\overline{k}\rangle\langle \overline{k}|A|\overline{l}\rangle\langle \overline{l}|m\rangle \\ &= \sum_{\overline{k},\overline{l}} S_{n\overline{k}} \left(A_P\right)_{\overline{k}\overline{l}} S^*_{m\overline{l}} = \sum_{\overline{k},\overline{l}} S_{n\overline{k}} \left(A_P\right)_{\overline{k}\overline{l}} S^\dagger_{\overline{l}m}\end{aligned}$$

or symbollically

$$A_Q = SA_PS^\dagger = T^\dagger A_P T \text{ or inversely}$$

$$A_P = TA_QT^\dagger.$$

Thus we see that when we use a transformation which takes us from one basis representation to another basis representation, all the formal equations between bras, kets, linear operators, etc. are the same as those when we consider a matrix transformation which referred to a given representation.

In the above treatment, we restricted ourselves to discrete indices for the eigenkets. But this need not be necessaritly so. One observable could have discrete indices depicting the eigenkets while the other observable could have continuouys indices depicting its eigenkets (Messiah Ch VII pp 21 discusses this case) or both observables could have continuous indices.

Because T and/or S are defined with respect to two representations, it is not easy to write down a formal representation of them as operators. However, there is one special case which merits mentioning. Suppose the Q-eigenkets and the P-eigenkets have a one-to-one correspondence. In this case we could "specialize" the relation

$$|u\rangle_Q = S|u\rangle_P \text{ to}$$
$$|n\rangle = S|overlinen\rangle$$

where the eigenkets $|n\rangle$ corresponds to the eigenkets $|\overline{n}\rangle$ in a one-to-one manner. Then we can write

$$S = S\sum_{\overline{m}} |\overline{m}\rangle\langle\overline{m}| = \sum_{\overline{m}} |\overline{m}\rangle\langle\overline{m}|$$

where $\sum_{\overline{m}}$ implies that m is summed over at the same time. In this case, S "straddles" or operates in both representations,

It is quite easy to show the scalar product is invariant to a change in the representation and that an observable A has the same eigenvalue spectrum.

6 Physical Framework of Quantum Mechanics

We have now discussed the more important mathematical aspects which we need to formulate quantum mechanics. Now we will give a set of postulates of quantum mechanics. This set of postulates are probably not complete. In fact, one may not be able to write down a set of postulates which covers everything that one might encounter in non-relativistic quantum mechanics. But these postulates will go "a long ways" towards setting up a framework to work in. In fact, I do not know of a textbook which lists a set of postulates and continues from there. Messiah in Chapter VIII does the most complete job in this respect I strongly recommend that you read this chapter in his book. A word of caution however, when reading various books, one must be careful to distinguish between postulates and definitions.

As you know, a classical system is described by a set of dynamical variables, e.q., position, momentum, energy, angular momentum. If these dynamical variables are known at any given instant of time, then we can in principle determine them at any other instant of time with infinite precision.

I repeat here the definition of an observable which we gave on page 48.

Definition A Hermitian operator is called an observable if its eigenkets form a complete set and span the entire vector space.

As I have pointed out before, it is usually not trivial to prove that a given set is complete. Consequently it is "traditional" for physicist to ignore this and to assume without proof that all Hermitian operators which correspond to physical quantities are observables. Part of this is contained in Postulate 2 below. But I have never seen anyone prove that the corresponding set pf eigenkets is indeed complete. The 1st postulate is:

Postulate 1. To every type of physical system, there corresponds an abstract Hilbert space and each vector in this space represents exactly one possible state of the system.

This postulate tells us that the system is described by a vector in a complete unitary linear vector space and so the various states of the system linearly "superposable". In addition, this "state vector" describing the system need not be constant but can move about in our Hilbert space as the system evolves in time – See Postulate 5 below.

In classical physics, the dynamical state of a system is described by a point in phase space and conversely each point in phase space represents a different state of the system. The same is also true for our Hilbert space with the exception that two vectors which are multiples of each other represent the same state.

The second postulate is:

Postulate 2. To each dynamical variable we associate a Hermitian linear operator which is an observable.

Now we have seen that not all operators in quantum mechanics will commute with each other. Those operators which do commute are said "to be compatible". If for a given system, we found all the compatible dynamical variables, we would say that we had a complete set of compatible variables. The precise determination of these variables of this complete set will give us the maximum information that we can obtain for this given quantum mechanical system. – Not that this is less information than one can obtain for a classical system since there one can determine all variables precisely and not just the compatible ones.

The 3rd postulate is:

Postulate 3. To any dynamical variable of a system, one can always add a certain number of other dynamical variables and thus form a complete set of compatible variables.

In choosing this complete set of compatible variable, one is usually "guided" by the dynamical

variables of the corresponding classical system. Does one know when he has them all? Generally not for a reasonably complicated system. Sometimes a variable in a quantum mechanical system does not have a classical analogue. – For example the intrinsic spin of a particle.

Disagreement with experimental results is one way one can tell that our set is incomplete. Just how one discovers and/or decides on what the "missing" variables are is basically a matter of intuition and/or insight and/or educated guesswork, etc.

Remember that the uncertainty principle tells us that non-compatible variables cannot be precisely determined at the same time:

THe 4th postulate is:

Postulate 4. If the system is in a state described by $|u\rangle$, the mean value of a physical quantity corresponding to the operator A is given by

$$\langle A\rangle = \frac{\langle u|A|u\rangle}{\langle u|u\rangle}.$$

This postulate tells us that we cannot tell the difference between two state vectors which are multiples of each other. In addition, it tells us that all states are undetermined up to an arbitray phase.

The 5th and final postulate is:

Postulate 5. Let $|\Psi(t)\rangle$ describe the state of the system at time t and let H be the Hamiltonian operator of the system. The system will evolve in time according to the equation

$$H|Psi(t)\rangle = i\hbar\frac{d}{dt}|Psi(t)\rangle.$$

Because this equation above is of first order in time, the linear superposition of states is preserved in the course of time. Messiah shows in Chapter 8, paragraph 8 that this equation follows quite naturally from the conservation of linear superposition of states if the system is conservative – i.e. H is not an explicity function of the time. However, we postulate that this equation is still valid even though H may be time dependent.

Let us now make some additional remarks concerning this complete ste of compatible observables which we need.

1. If the system has a classical analogue, we can take for the "fundamental observables" the

$3N$ coordinates q_i and the $3N$ corresponding conjugate momenta, p_i, along with the commutation relations:

$$[q_i, q_j] = 0, \quad [p_i, p_j] = 0, \text{ and } [q_i, p_j] = i\hbar\delta_{ij}.$$

All other observables, even H, can be written in terms of the p's and q's providing one properly symmetrizes "ambiguos" operators

$$\text{e.g. } p_i q_i \Rightarrow \frac{p_i q_i + q_i p_i}{2}.$$

2. If the system does not have a classical analogue, one usually has to introduce additional variables to those above and to specify the commutation relations among themselves and with the other observables which have a classical analogue. As I have pointed out above, finding these additional variables is not always easy and one does not know if he has them all. For example, the ground state of the hydrogen atom may be degenerate and we don't know it.

Let us suppose that we have somehow found this complete set of compatible variables, $A, B, C, \ldots$. Then in principle, we can find eigenstates $|a, b, c, \ldots\rangle$ which are simultaneous eigenkets of these compatible variables and which form a basis for our Hilbert Space.

Now our Hilbert Spaces can vary greatly depending on the system we are describing. As an example, consider a system containing a single particle with spin of $\hbar/2$ which can exist in only one of two states, one state with "spin up" and the other state with "spin down". In this case, our Hilbert Spaces is a 2-dimensional space with basis vectors $|+\hbar/2\rangle$ and $|-\hbar/2\rangle$.

On the other hand, consider a particle moving in a 1-dimensional spaces. Here the basis could be made of up the eigenkets, $|q\rangle$, of the position operator (or of the eigenkets, $|p\rangle$, of the momentum operator). For this simple system, out Hilbert Space is an infinite-dimensional space.

We have seen that when our eigenkets can be labeled by discrete indices

$$\langle n_i | n_j \rangle = \delta_{ij}$$

and that when they are labeled y continuous indices,

$$\langle q | q' \rangle = \delta(q - q').$$

Now it may seem odd to have a scalar product come out as a delta-function.

When we make the transition from our abstract Hilbert space to the coordinate representation (the Schrödinger Picture), we shall see that these δ-functions will always appear under the integral signs and thus will "give us no trouble".

Before we can make this transition to the coordinate representation, let us show that

$$\langle q'|p|q''\rangle = \frac{\hbar}{i}\delta'(q-q'') \text{ where } \delta'(x) = \frac{d}{dx}\delta(x).$$

In the coordinate representation, $|q'\rangle$ is an eigenket of the operator q with an eigenvalue of q', i.e., $q|q'\rangle = q'|q'\rangle$.

We now consider the Unitary Operator $S(\epsilon) = \exp^{-\frac{iq\epsilon}{\hbar}}$. Using $[q,p] = i\hbar$, it is quite easy to show that

$$[q, S(\epsilon)] = \epsilon S(\epsilon).$$

So $qS(\epsilon) = S(\epsilon)q + \epsilon S(\epsilon) = S(\epsilon)(q+\epsilon)$. Then we see that

$$q[S(\epsilon)|q'\rangle] = S(\epsilon)(q+\epsilon)|q'\rangle = S(\epsilon)(q'+\epsilon)|q'\rangle = (q'+\epsilon)[S(\epsilon|q'\rangle].$$

So $S(\epsilon)|q'\rangle = |q'+\epsilon\rangle$ and this $\langle q'|S)\epsilon|q''\rangle = \langle q'|q''+\epsilon\rangle = \delta(q'-q''-\epsilon)$. When ϵ is an infinitetesimal quantity, $S(\epsilon) \approx 1 - \frac{ip\epsilon}{\hbar}$ and then

$$\langle q'|S(\epsilon)|q''\rangle \approx \langle q'|1 - \frac{ip\epsilon}{\hbar}|q''\rangle = \langle q'|q''\rangle = \frac{i\epsilon}{\hbar}\langle q'|p|q''\rangle$$

or that

$$\langle q'|p|q''\rangle \approx \frac{\hbar}{i}\left[\frac{\delta(q'-q'') - \delta(q'-q''-\epsilon)}{\epsilon}\right]$$

and in the limit as $\epsilon \to 0$, we have

$$\langle q'|p|q''\rangle = \lim_{\epsilon\to 0}\frac{\hbar}{i}\left[\frac{\delta(q'-q'') - \delta(q'-q''-\epsilon)}{\epsilon}\right] = \frac{\hbar}{i}\delta'(q'-q'').$$

In a similar manner, one can show

$$\langle p'|q|p''\rangle = i\hbar\delta'(p'-p'').$$

6.1 The Schrödinger Picture or Wave Mechanics in the Coordinate Representation

What we will now show is that the Schrödinger Picture is nothing more than our quantum mechanics in our abstract Hilbert space in the coordinate representation. To do this, will consider a system which has a classical analogue with the 3 fundamental observables being the $3Nq_i$'s and the $3Np_i$'s where

$$[q_i, q_j] = 0, \quad [p_i, p_j] = 0, \text{ and } [q_i, p_j] = i\hbar\delta_{ij}.$$

The coordinate representation consists of picking as a basis the eigenkets of the $3Nq_i$ operators, i.e,

$$q_i'|q_1'.q_2', q_3', \ldots\rangle = q_i'|q_1'.q_2', q_3', \ldots\rangle.$$

Let us use the following notation:

$$\begin{aligned} |q'\rangle &= |q'_1, q'_2, q'_3, \ldots\rangle \\ dq' &= dq'_1\, dq'_2\, dq'_3\, \ldots \\ \delta(q'-q'') &= \delta(q'_1 = q''_1)\delta(q'_2 = q''_2)\delta(q'_3 = q''_3) \end{aligned}$$

Then our closure relation is given by

$$P = \int |q'\rangle\, dq'\, \langle q'| = I.$$

If $|\psi\rangle$ is a state vector in our Hilbert space describing a given state of the system, we can write using P

$$|\psi\rangle = \int |q'\rangle\, dq' \langle q'|\psi\rangle.$$

We have seen in the matrix representation that $\langle q'|\psi\rangle$ are elements of a column matrix for $|\psi\rangle$. In this case, the elements of this column matrix vary with a continuous index q'. Since $\langle q'|\psi\rangle$ is a scalar which varies continuously with q' and "anticipating what is to come", we define $\psi(q') \equiv \langle q'|\psi\rangle$ – i.e., $\psi(q')$ is a function of q'. Remember that our condensed notation real means

$$\langle q'|\psi\rangle = \langle q'_1 q'_2 q'_3 \ldots|\psi = \psi(q') = \psi(q'_1, q'_2, q'_3, \ldots).$$

Now let us look at the scalar product $\langle\psi_1|\psi_2\rangle$.

$$\begin{aligned} \langle\psi_1|\psi_2\rangle &= \left[\int \langle\psi_1|q'\rangle\langle q'|\, dq'\right]\left[\int |q''\rangle\langle q''|\psi_2\rangle\, dq''\right] \\ &= \int\int \langle\psi_1|q'\rangle\langle q'|q''\rangle\langle q''|\psi_2\rangle\, dq'\, dq'' \\ &= \int\int \langle\psi_1|q'\rangle\delta(q'-q'')\langle q''|\psi_2\rangle\, dq'\, dq'' \\ &= \int \langle\psi_1|q'\rangle\langle q'|\psi_1\rangle\, dq' = \int \psi_1^*(q')\psi_2^*(q')\, dq' \end{aligned}$$

Thus we see that in the coordinate representation, the scalar product $\langle\psi_1|\psi_2\rangle$ is the same scalar product that we would use in the Schrödinger Picture.

Next we look at some operator relationships involving the q_i's. The simplest is with q_i operating on $|\psi\rangle$. Consider $|\chi\rangle \equiv q_i|\psi\rangle$.

$$|\chi\rangle = q_i|psi\rangle = \int |q'\rangle\, dq'\, \langle q'|q_i|\psi\rangle = \int |q'\rangle\, dq'\, q'\langle q'|\psi\rangle$$

where we let the operator q_i operator to the left which we can do as it is a Hermitian operator. The components of $|\chi\rangle$ are

$$\begin{aligned} \chi(q'') = \langle q''|\chi\rangle &= \int \langle q''|q'\rangle\, dq'\, \langle q'|\psi\rangle = \int \delta(q''-q')\, dq'\, q'_i\langle q'|\psi\rangle \\ &= q''_i\psi(q''). \end{aligned}$$

Similarly, for any Hermitian operator which is an arbitrary function of the q_i's, we have for $V(q)$ that $\langle q'|V(q)|\psi\rangle = V(q')\langle q'|\psi\rangle = V(q')\psi(q')$ and this is just what we have or "use" in the Schrödinger Picture.

Now let us see what happens when one of the p_i's is our operator. Let $|\chi\rangle \equiv p_i|\psi\rangle$. THen

$$\langle q'|\chi\rangle = \langle q'|p_i|\psi\rangle = \int \langle q'|p_i|q''\rangle\langle q''|\psi\rangle\, dq''.$$

We saw on page page 64 that

$$\langle q_i'|p_i|q_i''\rangle = \frac{\hbar}{i}\delta'(q_i' - q_I'') = \lim_{\epsilon\to 0}\frac{\hbar}{i}\left[\frac{\delta(q'-q'') - \delta(q'-q''-\epsilon)}{\epsilon}\right].$$

Because of our condensed notation, we have here

$$\langle q'|p_i|q''\rangle = \frac{\hbar}{i}\delta'(q_1'-q_1'')\delta'(q_2'-q_2'')\ldots\delta'(q_{i-1}'-q_{i-1}'')\delta'(q_i'-q_i'')\delta'(q_{i+1}'-q_{i+1}'')\ldots.$$

and we can easily integrate over all the $dq'' = dq_1''dq_2''\ldots$ except the q_i'' one. With $\langle q''|\psi\rangle \equiv \psi(q'')$, we have

$$\begin{aligned}\langle q'|\chi\rangle &= \lim_{\epsilon\to 0}\frac{\hbar}{i}\left[\frac{\int \psi(q'')\delta(q_i'-q_i'')\,dq_i'' - \psi(q'')\delta(q_i'-q_i'-\epsilon)\,dq_i''}{\epsilon}\right] \\ &= \lim_{\epsilon\to 0}\frac{\hbar}{i}\left[\frac{\int \psi(q_1',q_2',q_3',\ldots,q_i'\ldots) - \psi(q_1',q_2',q_3',\ldots,q_i'-\epsilon\ldots)}{\epsilon}\right] = \frac{\hbar}{i}\frac{\partial}{\partial q_i'}\psi(q')\end{aligned}$$

So we see that

$$\langle q'|\chi\rangle = \langle q'|p_i|\psi\rangle = \frac{\hbar}{i}\frac{\partial}{\partial q_i'}\psi(q').$$

and this is just how the operator p_i is "supposed to behave" in the Schrödinger Picture.

Clearly we can generalize these results so that for any operator $A(p,q)$ operating on $|\psi\rangle$ will be represented in the Schrödinger Picture by $A\left(\frac{\hbar}{i}\frac{\partial}{\partial q}, q\right)$ operating on $\psi(q)$.

Furthermore, the equation

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

clearly can be "replaced" by

$$H\psi(t) = i\hbar\frac{\partial}{\partial t}\psi(t)$$

where $\frac{d}{dt} \to \frac{\partial}{\partial t}$ because the eigenvalues of the operator q appear explicitly in the component $\psi(q,t)$.

So we are now back to the Schrödinger Picture. But we should stress that this is just quantum mechanics in a given representation – the coordinate representation. This is just one of the many valid representations which one could work in. Another common representation is the momentum

representation. Here the operator p_i is replaced by its eigenvalue p_i and the operator q_i is replaced by $-\frac{\hbar}{i}\frac{\partial}{\partial p_i}$ where it operators on $\langle p'|\psi\rangle \equiv \phi(p')$.

In principle, one need not even go into a representation to find out all there is to know about a quantum mechanical system as I will show shortly for the simple case of the linear harmonic oscillator.

One might say that we are now in the same position as a student studying E & M who has been given Maxwell's Equations – all there is left to do is to apply them to various systems with various boundary conditions. So for the rest of the year expect for relativistic quantum mechancs and the Dirac equation, we will look at various systems and learn techniques on how to "solve" them and when to use different approximation.

Usually we will work in the Schrödiinger Picture but we need not to do so.

7 One Dimensional Harmonic Oscillator

7.1 1. Review

First let us review the standard treatment of the linear harmonic oscillator in the coordinate representation which you can find in practically all text books of Quantum Mechanics, e.g, Pauling & Wislon or Schiff pp. 67 ff. I will assume that you have already done this in a more elementary course in Quantum Mechanics and so I'll only review the more important aspects here.

The harmonic oscillator or something closely related to it is forever appearing the physics of all types of phenomena – mechanics, E & M, solid state, etc. – and is often used as the starting point to solving more complicated sytems.

Our Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \text{ where } \omega^2 \equiv \frac{k}{m}.$$

Since HH is time independent, we can write $\Psi(x,t) = \psi(x)e^{-\frac{iEt}{\hbar}}$ and our wave equation

$$H\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t)$$

reduces to

$$H\psi(x) = E\psi(x) \text{ or } \left(\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + E - \frac{1}{2}m\omega^2 x^2\right)\psi(x) = 0.$$

At this point, it is convenient to change variables and we define

$$E = \lambda\left(\frac{\hbar\omega}{2}\right) \text{ and } x = \xi\sqrt{\frac{\hbar}{m\omega}}.$$

Our differential equations becomes

$$\frac{d^2\psi(\xi)}{d\xi^2} + (\lambda - \xi^2)\psi(\xi) = 0.$$

For large ξ we can neglect λ and we see that our solution goes like $e^{-\frac{\xi^2}{2}}$. So we look for a solution of the form

$$\psi(\xi) = \phi(\xi)\exp^{-\frac{\xi^2}{2}}$$

which gives us a differential equation of

$$\frac{d^2\phi}{d\xi^2} - 2\xi\frac{d\phi}{d\xi} + (\lambda - 1)\phi = 0.$$

Now a standard approach is to try a series solution of the form

$$\psi(\xi) = \xi^s(b_o + b_1\xi + b_2\xi^2\ldots) \text{ where } b_0 \neq 0.$$

Substituting this solution into the differential equation we find by equating each coefficient of each power of ξ to zero that $s = 0$ or 1 from the 1st equation and so $s = 0$ and/or $b_1 = 0$ from the 2nd equation. The succeeding equations five use recursion relations so that we have

$$\begin{aligned}
\phi(\xi) &= a_0 + a_2\xi^2 + a_4\xi^4 + \ldots \text{ for } s = 0 \text{ and} \\
\phi(\xi) &= a_1\xi + a_3\xi^3 + a_5\xi^5 + \ldots \text{ for } s = 1 \text{ with} \\
a_{k+2} &= \frac{2k - \lambda + 1}{(k+1)(k+2)}a_k.
\end{aligned}$$

One can quite easily show that both solutions $\psi(\xi) = \exp^{-\frac{lxi^2}{2}}\phi(xi)$ for $s = 0$ or 1 diverge as $x \to \pm\infty$.But this is not acceptable as $\psi(x)$ must remain finite. $\psi(x)$ can only remain finite if $\phi(\xi)$ is a finite series which will occur if λ is restricted to only take on the values $2n + 1$ where $n = 0, 1, 2, \ldots$. This restricts the possible values of E and we find the quantized energy eigenvalues of

$$E_n = (n + \frac{1}{2})\hbar\omega.$$

By picking the two arbitrary constants a_0 and a_1 "correctly", the $phi(\xi)$'s become the Hermite polynomials:

$$\begin{aligned}
H_0(\xi) &= 1 \\
H_1(\xi) &= 2\xi \\
H_2(\xi) &= 4\xi^2 - 2i \\
H_3(\xi) &= 8\xi^3 - 12\xi 2i \\
\ldots &\quad \ldots \\
H_n(\xi) &= (-1)^n \exp^{\frac{\xi^2}{2}} \frac{d^n e^{-\xi^2}}{d\xi^n}.
\end{aligned}$$

$H_n(\xi)$ satisfies the differential equation of $\phi(\xi)$ with $\lambda = 2n+1$, viz.

$$H_n'' - 2\xi H_n' + 2nH_n = 0.$$

One can also show that

$$\begin{aligned} H_n' &= 2nH_{n-1} \\ 0 &= H_{n+1} - 2\xi H_n + 2nH_{n-1} \\ H_{n+1} &= 2\xi H_n - H_n' \end{aligned}$$

In fact, one can take this last equation along with $H_0(\xi) = 1$ and "generate" all the other H_n's.

One can also use a generating function to obtain the H_n's. The generating function $g(\xi, h)$ is given by

$$g(\xi, h) = e^{2\xi h - h^2} = \sum_{n=0}^{\infty} \frac{H_n(\xi) h^n}{n!}.$$

The H_n's are obtained by expanding $e^{2\xi h - h^2}$ and equating coefficients of like powers of h on both sides. These generating functions are vey useful in evaluating integrals involving the H_n's.

One usually normalizes $\psi(x)$ to unity by

$$|n\rangle \equiv \psi_n(\xi) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} e^{-\frac{\xi^2}{2}} H_n(\xi)$$

and then

$$\int \psi_n^*(x)\psi_m(x)\,dx = \sqrt{\frac{\hbar}{m\omega}} \int \psi_n^*(\xi)\psi_m(\xi)\,d\xi = \delta_{nm} \text{ and}$$

$$H\psi_n(\xi) = E_n \psi_n(\xi).$$

The matrix elements $\langle n|H|m\rangle$ form a diagonal matrix in the coordinate representation

$$H = \begin{pmatrix} \frac{1}{2}\hbar\omega & 0 & 0 & 0 & 0 & \dots \\ 0 & \frac{3}{2}\hbar\omega & 0 & 0 & 0 & \dots \\ 0 & 0 & \frac{3}{2}\hbar\omega & 0 & 0 & \dots \\ 0 & 0 & 0 & \frac{7}{2}\hbar\omega & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

and one can show that

$$\xi = \begin{pmatrix} 0 & \sqrt{\frac{1}{2}} & 0 & 0 & 0 & 0 & \dots \\ \sqrt{\frac{1}{2}} & 0 & \sqrt{\frac{2}{2}} & 0 & 0 & 0 & \dots \\ 0 & \sqrt{\frac{2}{2}} & 0 & \sqrt{\frac{3}{2}} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{\frac{3}{2}} & 0 & \sqrt{\frac{4}{2}} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

7.2 2. An Alternatice Approach

After this quick review of the harmonic oscillator, I now want to show you what I promised earlier. That is, given the fundamental observables of the system and their commutation relations, I can solve this system without ever going into a representation. In fact, one should always be able to do this in principle but it is generally not easy to do so in practive.

In this case, our fundamental observables are p and q (I will now use q instead of x) and we have

$$[p,p] = [q,q] = 0 \text{ and } [q,p] = i\hbar.$$

All our observables and/or operators are functions of p and q. In particular

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 = .$$

Instead of using p and q, it is more conventient to introduce two other operators $R^\dagger$ and R where

$$\begin{aligned} R^\dagger &\equiv \frac{1}{\sqrt{2m\omega\hbar}}(m\omega q - ip) \qquad \text{and} \\ R &\equiv \frac{1}{\sqrt{2m\omega\hbar}}(m\omega q + ip) \end{aligned}$$

or inversely

$$\begin{aligned} q &= \sqrt{\frac{\hbar}{2m\omega}}(R^\dagger + R) \qquad \text{and} \\ p &= \sqrt{\frac{m\omega\hbar}{2}}(R^\dagger - R) \end{aligned}$$

Using $[q,p] = i\hbar$, we find that

$$[R, R^\dagger] = 1, \quad [R^\dagger, R^\dagger] = 0, \qquad \text{and } [R,R] = 0.$$

For H, we find

$$H = (R^\dagger R + \frac{1}{2})\hbar\omega = (R^\dagger R + RR^\dagger)\frac{\hbar\omega}{2}.$$

So our eigenvalue equaton

$$H|\Psi\rangle = E|\Psi\rangle$$

reduces to solving an eigenvalue equation for $R^\dagger R$. That is, if we can solve

$$R^\dagger R|n\rangle = \lambda_n|n\rangle$$

then due to the form of H, we can certainly find E_n in

$$H|n\rangle = E_n|n\rangle.$$

First we show that λ_n is non-negative. For this, consider $|\chi\rangle \equiv R|n\rangle$. Then

$$\langle\chi|\chi\rangle = \langle m|R^\dagger R|n\rangle = \lambda_n\langle n|n\rangle$$

and since $\langle\chi|\chi\rangle$ and $\langle n|n\rangle$ are positive definite, we need

$$\lambda_n \geq 0.$$

Next we look at

$$|\chi\rangle \equiv R^\dagger|n\rangle$$

and operate on it with $R^\dagger R$. Then

$$\begin{aligned} R^\dagger R|\chi\rangle &= R^\dagger RR^\dagger|n\rangle = R^\dagger(1 + R^d agger R)|n\rangle = R^\dagger(1+\lambda_n)]ketn \\ &= (\lambda_n+1)(R^\dagger|n\rangle) = (\lambda_n+1)|\chi\rangle. \end{aligned}$$

So we see that $|\chi\rangle = R^\dagger|n\rangle$ is an eigenket of $R^\dagger R$ with an eigenvalue $(\lambda_n + 1)$. Likewise $R^\dagger R^\dagger|n\rangle$ is an eigenket with eigenvalue $(\lambda_n + 2)$. Thus, given $|n\rangle$, we can "generate" all the higher eigenkets of $R^\dagger R$.

In a like manner, we can show that $R|n\rangle$ is also an eigenfunction of $R^\dagger R$ with an eigenvalue of $\lambda_n - 1$...and so forth for $RR|n\rangle$, $RRR|n\rangle, \ldots$. While we could increase λ_n indefinitely by using $R^\dagger$'s, we cannot reduce it indefinitely by using R's as we must satisfy $\lambda_n \geq 0$ and so we cannot have negative eigenvalues. If we operate on each successively smaller eigenket with R, we will eventually reach the eigenket with the "lowest possible" eigenvalue. Let's call this eigenket $|\lambda_0\rangle$. Then upon operating with R again, we must have

$$R|\lambda_0\rangle = 0$$

or otherwise we will generate an eigenket with a lower eigenvalue which will be a contradiction. Now λ_0 is somewhere between zero and one. We have

$$R^\dagger R|\lambda_0\rangle = \lambda_0|\lambda_0\rangle \qquad \text{and } R|\lambda_0\rangle = 0.$$

But $R^\dagger R|\lambda_0\rangle = R^\dagger(R|\lambda_0\rangle) = R^\dagger(0) = 0$ which tells us that $\lambda_0 = 0$ uniquely.

Instead of labeling our eigenkets by $|\lambda_0\rangle, |\lambda_1\rangle, |\lambda_2\rangle, \ldots$, let us use $|0\rangle, |1\rangle, |2\rangle, \ldots, |n\rangle, \ldots$. For these eigenkets, we have

$$R^\dagger R|n\rangle = n|n\rangle$$

with $n = 0, 1, 2, \ldots$ and n has no upper limit.

Clearly since $H = (R^\dagger R + \frac{1}{2})\hbar\omega$, we have

$$H|n\rangle = (n + \frac{1}{2})\hbar\omega = E_n|n\rangle.$$

At this point, we have found the eigenspectrum of the Hamiltonian H which has eigenvalues of $\frac{1}{2}\hbar\omega, \frac{3}{2}\hbar\omega, \frac{5}{2}\hbar\omega, \ldots$ and we did it without going into an explicit representation – That is, we are still working on our abstract Hilbert space. It is also clear that H is diagonal in terms of these eigenkets.

To complete the picture, all we have to do now is demonstrate how one obtains or calculates matrix elements of p and q or or arbitrary functions of p and q.

First we make $R^\dagger|n\rangle \propto |n+1\rangle$ a normalized eigenket. Let $|\chi\rangle \equiv NR^\dagger|n\rangle$ and we want $\langle\chi|\chi\rangle = 1$ along with $\langle n|n\rangle = 1$. So we have

$$\langle\chi|\chi\rangle = |N|^2\langle n|RR^\dagger|n\rangle = |N|^2\langle n|1 + R^\dagger R|n\rangle = |N|^2(n+1)\langle n|n\rangle = 1.$$

So we pick (up to an arbitrary phase) $N = \dfrac{1}{\sqrt{n+1}}$. This

$$\begin{array}{lrcl} \text{Thus} & |n+1\rangle & = & \frac{1}{\sqrt{n+1}}R^\dagger|n\rangle \\ \text{Likewise} & |n-1\rangle & = & \frac{1}{\sqrt{n}}R|n\rangle \end{array} \quad \left\{ \begin{array}{c} \text{“Rule” – Use square} \\ \text{root of largest} \\ \text{eigenvalue appearing} \end{array} \right.$$

Now we are ready to look at the matrix elements of $\xi = q\sqrt{\dfrac{m\omega}{\hbar}}$ – that is, the matrix elements of q. As $\xi = \dfrac{1}{\sqrt{2}}(R^\dagger R)$, we see that

$$\langle n|\xi|m\rangle = \frac{1}{\sqrt{2}}\langle n|R^\dagger R|m\rangle = \sqrt{\frac{m+1}{2}}\langle n|m+1\rangle + \sqrt{\frac{m}{2}}\langle n|m-1\rangle$$

or

$$\langle n|\xi|m\rangle = \sqrt{\frac{m+1}{2}}\delta_{n,m+1} + \sqrt{\frac{m}{2}}\delta_{n,m-1}.$$

Really easy!! Compare with ξ-matrix on page page 69. For the homework, we will evaluate $\langle n|\xi|m\rangle$ using generating functions and you will appreciate how easy this method is.

Clearly we can continue and obtain the matrix elements of q^2, p, or any combinations of p's and q's using this approach.

Thus we have solved the harmonic oscillator without going into an explicit representation. If one wishes, he can start at this point and “go into” the coordinate representation. For details see Schiff pages 184 and 185.

7.3 3. A Different Point of View

We now look at this new approach of using the $R^\dagger$'s and R's in a different manner. Suppose we consider some elementary excitations which we call “particles” or “quanta” and which have an

energy of $\hbar\omega$. Then we could say that the state $|n+1\rangle$ had one more of these "particles" in it than the state $|n\rangle$. Similarly the state $|n-1\rangle$ has one "particle" less. Taking this point of view, $R^\dagger$ would when acting on a state create an additional one of these particles and R would annihilate or destroy one of these particles. That is $R^\dagger$ and R are creation and destruction operators for these particles respectively. In this interpretation $R^\dagger R$ is the counting operator which "counts" the number of paricles in a given state.

These creation and destruction operators obey the commutation relation

$$[R^\dagger.R^\dagger] = 0, \quad [R, R] = 0, \text{ and } [R, R^\dagger] = 1.$$

Particles which do this are called bosons and there wave functions are symmetric under interchange of any two particles. Furthermore, there is no restriction on how many one can put into the same state with an energy $\hbar\omega$.

To see this symmetry under interchange of any two particles, consider the state

$$|n\rangle = \frac{1}{\sqrt{n!}} R^\dagger R^\dagger \dots R^\dagger |0\rangle = \frac{1}{\sqrt{n!}} (R^\dagger)^n |0\rangle.$$

Perhaps is is clearer to put a subscript labelling each particle on the $R^\dagger$'s. Then

$$|n\rangle = |123\dots n\rangle = \frac{1}{\sqrt{n!}} R_n^\dagger R_{n-1}^\dagger \dots R_3^\dagger R_2^\dagger R_1^\dagger |0\rangle.$$

Now we see that interchanging any two particles will give the same wave function back since the $R^\dagger$'s commute.

An obvious generalization of the above approach is to have many linear oscillators all of different frequencies labeld by ω_k and with respective creation and annihilarion operators $R_k^\dagger$ and R_k. Our commutation relations are

$$[R_k, R_l] = 0, [R_k^\dagger, R_l^\dagger] = 0 \qquad \text{and } [R_k, R_l^\dagger] = \delta_{kl}.$$

Lager on in the year we will be studying radiation from atoms. There we will see that the electromagnetic field can be represented by an infinite set of harmonic oscillators of all frequencies. We will then "quantize" these oscillators and the bosons resulting from this quantization will be the quanta of the electromagnetic field – Photons.

If our creation and annhilation operators had obeyed anticommutation relations instead of commutation relations, we would have a system of fermions instead of bosons. Then one can show that the wave function is antisymmetric on interchange of any two particles and, at most, only one particle can occupy a given state at a time.

When one describes a system by using creation and annihiltion operators for the particles of the system, one says that he is using the "occupation number representation" or the "formulation of second quantization".

A glimpse of "bigger and better" things? If we can quantize the electromagnetic field, why can't we quantize the field describing the wave function $\psi(\vec{r})$? People have tried this and they originally hoped that the quanta of these fields would somehow be connected to elementary particles. But when people tried to quantize the field $\psi(\vec{r})$, they ran into all kinds of difficulties – mostly of a mathematical nature, These problems have not been resolved and their treatment is beyond the scope of this course.

8 Central Potentials

Let us know look at the special case where we have spinless particles moving in the presence of a central potential. These central potentials will only depend on the magnitude of $\vec{r}$ (i.e., they are spherically symmetric) and will be denoted by $V(r)$. Introduction of intrinsic spin is straightforward by is an additional complication which does not give any additional physical insight into the problem – So we will delay its introduction until later. Furthermore we will only look at the bound states at this time.

For a single particle moving in a central potential, the time independent wave equation is

$$H\psi(\vec{r}) = \left(\frac{p^2}{2m} + V(r)\right)\psi(\vec{r}) = E\psi(\vec{r}) \quad \text{or}$$

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(r) - E\right)\psi(\vec{r}) = 0.$$

If we stay in the Cartesian coordinate system, one cannot go much further until $V(r)$ is specified. But because of the spherical symmetry of $V(r)$, it is much more desirale to approach this problem in spherical coordinates. To do this, we Introduce the orbital angular momentum operator $\vec{L} = \vec{r} \times \vec{p}$. In component form, we have $L_x = yp_z - zp_y$ (cyclic). By cyclic, we mean that this formula holds for $x \to y \to z \to x$. Also we have $[x, p_x] = i\hbar$ (cyclic). With these relations, it is easy to show that

$$[L_x, L_y] = i\hbar L_z \quad \text{(cyclic)}.$$

Symbolically, we can write the three equations implied above as

$$\vec{L} \times \vec{L} = i\hbar\vec{L}.$$

Using the above commutation relations for the L_i's, one can easily show with

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

that

$$[L_x, L^2] = [L_y, L^2] = [L_z, L^2] = 0.$$

In spherical coordinates, one has

$$\begin{aligned} L_x &= i\hbar\left[\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right], \\ L_y &= -i\hbar\left[\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right], \\ L_z &= -i\hbar\frac{\partial}{\partial\phi}, \end{aligned}$$

and

$$L^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right].$$

With this expression for L^2, one can easily show that in spherical coordinates,

$$\begin{aligned} H = \frac{p^2}{2m} + V(r) &= -\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{L^2}{2mr^2} + V(r) \\ &= -\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \frac{L^2}{2mr^2} + V(r) \end{aligned}$$

So we see that all the θ and ϕ dependence of our Hamiltonian operator (or of ∇^2) is contained in the operator L^2.

The standard approach to solving

$$H\psi(\vec{r}) = E\psi(\vec{r})$$

is to assume separable solutions, i.e,

$$\psi(\vec{r}) = R(r)\Theta(\theta)\Phi(\phi) = R(r)f(\theta,\phi).$$

I will assume that you have already done this in an elementary course in quantum mechanics. What you did and/or found was that one could solve for $f(\theta,\phi)$ which was a solution to the differential equation

$$L^2 f(\theta,\phi) = cf(\theta,\phi).$$

The solutions to this differential equation are the spherical harmonics $Y_{lm}(\theta,\phi)$ which consists of products of associated Legendre polynomials and $e^{\pm im\phi}$'s. The solutions, Y_{lm} have the properties of

$$\begin{aligned} L^2 Y_{lm}(\theta,\phi) &= l(l+1)\hbar^2\, Y_{lm}(\theta,\phi) \text{ and} \\ L_z Y_{lm}(\theta,\phi) &= m\hbar\, Y_{lm}(\theta,\phi) \end{aligned}$$

So the Y_{lm}'s are simultaneous eigenfunctions of L^2 and L_z. – This is okay as L^2 and L_z commute. We could have easily looked for simultaneous eigenfunctions of L^2 and L_x or of L^2 and L_y. – But we couldn't look for simultaneous eigenfunctions of L_y and L_z as they don't commute, etc. for L_x and

L_z or L_x and L_y. The eigenfunctions of L^2 and L_x or of L^2 and $_y$ can always be expanded as $\sum_m a_m Y_{lm}$ and these eigenfunctions will still be eigenfunctions of L^2. However, it is simpler and conventional to use the eigenfunctions of L^2 and L_z, viz., the Y_{lm}'s.

In quantum mechanics, we know that all eigenfunctions are undefined up to an arbitrary phase, Recent quantum mechanical texts usually adopt what is known as the "Condon and Shortley" phases for the Y_{lm}'s and we will do this also. A word of caution however, all texts and/or articles do not necessarily do so – see Pauling & Wilson, 1st edition of Schiff. The phases when we adopt fo the Y_{lm}'s are

$$\begin{aligned} Y^*_{l,m}(\theta,\phi) &= (-1)^m\, Y_{l,-m}(\theta,\phi) \qquad \text{with} \\ Y_{lm}(\theta,\phi) &= (-1)^m \sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}}\, P_l^{|m|}(\cos\theta) e^{im\phi}. \end{aligned}$$

The Y^*_{lm}'s and $Y_{lm}(\theta,\phi)$'s are orthonormal such that

$$\int Y^*_{lm}(\theta,\phi) Y_{l'm'}(\theta,\phi)\, \sin\theta\, d\theta d\phi \;= \int Y^*_{l'm'}(\Omega) Y_{l'm'}(\Omega)\, d\Omega = \delta_{ll'}\delta_{mm'}$$

where $d\Omega \equiv \sin\theta\, d\theta d\phi$ = differential element of solid angle. (See Schiff pp. 76-83 for more details).

With this digression dealing with the spherical harmonics, $Y_{lm}(\theta,\phi)$, which by the way form a complete set in θ,ϕ space, let us now return to our central potential problem with

$$H = \frac{p^2}{2m} = -\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2} + \frac{L^2}{2mr^2} + V(r).$$

Instead of solving $H\psi(\vec{r}) = E\psi(\vec{r})$ using separable solutions, we note that the $Y_{lm}(\theta,\phi)$'s form a complete set and so we expand $\psi(\vec{r})$ in terms of them, viz.,

$$\psi(\vec{r}) = \sum_{l',m'} R_{l'm'} Y_{l'm'}(\theta,\phi).$$

Since $L^2 Y_{l'm'} = l'(l'+1)\hbar^2 Y_{l'm'}$, our differential equation becomes

$$\sum_{l'm'} \left[\frac{1}{r}\frac{\partial^2}{\partial r^2} r - \frac{l'(l'+1)}{r^2} + \frac{2m}{\hbar^2}(E - V(r))\right] R_{l'm'} Y_{l'm'} = 0$$

Multiplying by $Y^*_{lm}(\theta,\phi)$ and integrating over θ and ϕ, the orthogonality of the $Y_{lm}(\theta,\phi)$'s give us

$$\left[\frac{1}{r}\frac{\partial^2}{\partial r^2} r - \frac{l(l+1)}{r^2} + \frac{2m}{\hbar^2}(E - V(r))\right] R_{lm}(r) = 0$$

which is the differential equation which we must solve for $R_{lm}(r)$, the radial solutions.

For a central potential, $V(r)$, the "magnetic quantum number", m, does not enter – the m in $\frac{2m}{\hbar^2}$ is the mass – and so we can drop the subscript m on $R_{lm}(r)$ and write it simple as $R_l(r)$.

In the more general case where $V(\vec{r})$ is not spherically symmetric, $R_{lm}(r)$ will depend on m and we will have to retain it in our solution.

If our particles have intrinsic spin, $\vec{(}s)$, the only change will be to replace $Y_{lm}(\theta,\phi)$ above by $Y_{lm}(\theta,\phi)\chi_{sm_s}$ where X_{sm_s} are eigenfunctions of the spin operator. – We will do this in more detail later.

In order to proceed further, we need to now the form of $V(r)$. We will now look at three rather common cases – The square well, the hydrogen atom, and the three dimensional harmonic oscillator. I will assume that you have already done these three cases before. I will only summarize them and let you fill in the details.

8.1 Square Well (Bound State

For a square well, we have

$$\begin{aligned} V(r) &= -V_0 \quad \text{for} \quad 0 \le r \le a \\ V(r) &= 0 \quad \text{for} \quad r > a \end{aligned}$$

We define

$$\begin{aligned} K &= \sqrt{\frac{2m}{\hbar^2})V_0 - |E|} \text{ and } \rho = Kr \text{ for } r \le a \\ k &= \sqrt{\frac{2m|E|}{\hbar^2}} \text{ and } \rho = ikr \text{ for } r > a \end{aligned}$$

For $r \le a$ and $\rho = Ka$, our differential equation becomes

$$\frac{d^2R_l(\rho)}{d\rho^2} + \frac{2}{\rho}\frac{dR_l(\rho)}{d\rho} + \left[1 - \frac{l(l+1)}{\rho^2}\right] R_l(\rho) = 0.$$

The solutions to this differential equation are the spherical Bessel functions $j_l(\rho)$ and $n_l(\rho)$ where

$$\begin{aligned} k_l(\rho) &= \sqrt{\frac{\pi}{2\rho}} J_{l+\frac{1}{2}}(\rho) \text{ and is regular at the origin and} \\ n_l(\rho) &= (-1)^{l+1}\sqrt{\frac{\pi}{2\rho}} J_{-(l+\frac{1}{2})}(\rho) \text{ and is irregular at the origin.} \end{aligned}$$

Since $n_l(\rho)$ is irregular at the origin, it must be excluded in the solution $R_l(\rho)$ for $r \le a$. Then $R_l(\rho) \propto j_l(\rho)$ for $r \le a$.

The asymptotic properties of j_l and n_l which we will use quite often are

$$\begin{aligned} \lim_{\rho\to 0} j_l(\rho) &\to \frac{\rho^l}{(2l+1)!!}, \qquad \lim_{\rho\to\infty} j_l(\rho) \to \frac{\cos[\rho-(l+1)\frac{\pi}{2}]}{\rho}, \\ \lim_{\rho\to 0} n_l(\rho) &\to \frac{(2l-1)!!}{\rho^{l+1}}, \qquad \lim_{\rho\to\infty} n_l(\rho) \to \frac{\sin[\rho-(l+1)\frac{\pi}{2}]}{\rho}. \end{aligned}$$

where $(2l+1)!! \equiv (2l+1)(2l-1)(2l-3)\ldots 5\cdot 3\cdot 1$.

In the "exterior" region where $r \geq a$, we have the same differential equation for $R_l(\rho)$ with $\rho = ikr$. Since the origin is now excluded we can expect that $R_l(\rho)$ will be in general a linear combination of j_l and n_l. Normally it is more convenient to use instead the spherical Hankel functions $h_l^{(1)}(\rho)$ and $h^{(2)}(\rho)$ where

$$\begin{aligned} h_l^{(1)}(\rho) &\equiv j_l(\rho) + in_l(\rho) \quad \text{and} \\ h_l^{(2)}(\rho) &\equiv j_l(\rho) - in_l(\rho), \end{aligned}$$

which have asymptotic behaviors of

$$\begin{aligned} \lim_{\rho\to\infty} h_l^{(1)}(\rho) &\to \frac{e^{i[\rho-(l+1)\frac{\pi}{2}]}}{\rho} \quad \text{and} \\ \lim_{\rho\to\infty} h_l^{(2)}(\rho) &\to \frac{e^{-i[\rho-(l+1)\frac{\pi}{2}]}}{\rho}. \end{aligned}$$

For our case of a square well, $\rho = -ikr$ and we see that the Hankel functions will describe increasing and decreasing exponentials in the asymptotic region where $rho \to \infty$. Clearly, the increasing exponential will not be allowed for the bound states.

Furthermore, our solutions, and their derivatives must be continuous across the "boundary" at $r = a$. Not every value of E will satisfy this "matching requirement". As a result, only certain discrete values of E will be allowed for bound states. Practically all quantum mechanics texts work out these solutions and I refer you to them for details.

For an infinitely deep square well, the energy levels look somewhat like those shown at the right Instead of labeling of the levels by $l = 0, 1, 2, 3, \ldots$, one uses $s, p, d, f, g, h, i, \ldots$ respectively. The lowest

Figure 1: Infinite Square Well Energy Eigenvalues

s levels is the $1s$ level, the next is the $2s$ level, etc.

This is one of the two most common labelling schemes. Another labeling scheme will be used for the Hydrogen atom which we will discuss next.

8.2 Hydrogen Atom (without spin)

Now we have a two-body problem but we will show that it reduces to a one-body problem. Our Hamiltonian is

$$H = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + V(|\vec{r_1} - \vec{r_2})|) \equiv E_{total}$$

where one particle is an electron, the other is a proton, and $V(|\vec{r_1} - \vec{r_2})|)$ is the Coulomb potential between them. We now transform to the relative and center of mass coordinate using

$$\vec{r} = \vec{r_1} - \vec{r_2}, \quad \vec{R} = \frac{m_1\vec{r_1} + m_2\vec{r_2}}{m_1 + m_2}, \quad M = m_1 + m_2, \text{ and}$$

$$m = \frac{m_1 m_2}{m_1 + m_2} \quad = \text{ reduced mass.}$$

This transformation gives

$$H = \frac{P^2}{2M} + \frac{p^2}{2m} + V(r), \quad \text{where} \quad \vec{P} = M\dot{\vec{R}}, \quad \text{and} \quad \vec{p} = m\dot{\vec{r}}.$$

We see that the Hamiltonian separates into two parts and so we can seek a separable solution

$$\psi(\vec{r_1}, \vec{r_r}) = \psi_{CM}(\vec{R})\psi(\vec{r}).$$

The differential equation for $\psi_{CM}(\vec{R})$ is

$$\frac{P^2}{2M}\psi_{CM}(\vec{R}) = E_{CM}\psi_{CM}(\vec{R})$$

where $E_{total} \equiv E_{CM} + E$.