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 Lithschitz — 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.

Index to 215A Notes

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 Schroedinger 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,

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

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

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

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

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

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

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

This can be written as

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

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

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

$\vec{S}$ 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.,

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

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.,

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

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 $f_1 = \int Gg_1\,dx$ and $f_2 = \int Gg_2\,dx$, 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 $\begin{Bmatrix}\text{positive}\\ \text{negative}\end{Bmatrix}$, the integral is still zero as $\lim_{N=\infty} NA = \lim_{N\to\infty} NB = \{\pm\infty\}$. 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. $\delta(ax) = \frac{1}{a}\delta(x)$ 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. $\delta(x^2 - a^2) = \frac{1}{2a}[\delta(x - a) + \delta(x + a)]$
8. $x\delta'(x) = x\frac{d}{dx}\delta(x) = -\delta(x)$
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 $\int_a^a f\,d\alpha = 0$.

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 $\displaystyle\int_a^b f\, d\alpha$ 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 $\displaystyle\int_a^b f\, d\alpha$ 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 $\begin{Bmatrix} \Theta(x) = 0 \text{ for } 0 < x \\ \Theta(x) = \text{‘ for } x > 0 \end{Bmatrix}$ 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.
Npw

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

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).
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 soread in coordinate space, Δx, will be large and conversely. This is result is also consistent with our uncertainty principle which was $\Delta x\Delta p \geq \hbar/2$ or $Deltax\Delta k \geq 1/2$.

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 <u>any</u> 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.

<u>Definition of a Field</u>

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. $\alpha\alpha^{-1} = I \text{ for} \alpha \neq 0$ — 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.

<u>Definition of a Linear Vector Space</u>

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 <u>linear</u> <u>vector</u> <u>space</u>, 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.*

Definition of 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.

Definition of 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.