Notes on Quantum States and Observables

October 11, 1995

Massachusetts Institute of Technology

Department of Physics

Physics 8.04 October 11, 1995

The first step in building a theory capable of predicting the results of experiments carried out on a physical system is to develop a description of the different possible physical *states* of the system and a method for the prediction of the outcome of experiments carried out on a system in a given state. Only once this is done do we then go about the business of building a theory to give us the possible states of a system and to describe how a system moves from state to state as a function of external and internal influences.

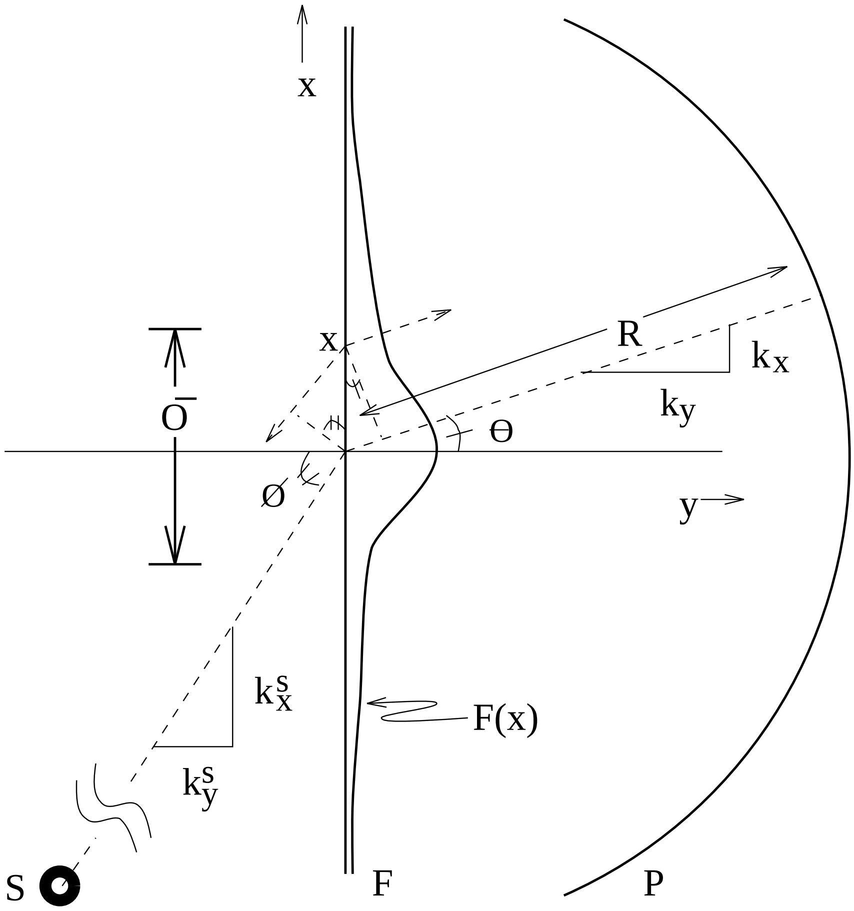
1 States and Observables

In classical mechanics, the state of a system may be specified by giving the coordinates $\{\vec{x}_i\}$ and velocities $\{\vec{v}_i\}$ of all of the particles in the system. Once this is done, all measurable quantities such as the momenta $\{\vec{p}_i\}$, total energy $H = \sum \frac{p_i^2}{2m} + V(\{x_i\})$ and angular momentum $\vec{L} = \sum \vec{r}_i \times \vec{p}_i$ of the system may be determined. Such measurable quantities are known as *observables.* Under the classical principle of determinism, once the state of the system is specified at time $t = 0$, the laws of motion may be solved to determine the state of the system $\{\vec{x}_i(t), \vec{v}_i(t)\}$ for all later times.

The situation in quantum mechanics is not much different. We may still make measurements of the familiar physical observables mentioned above (and some new ones such as spin or polarization which we will learn about later). As we have seen in our discussions of the uncertainty principle and interference experiments, although the results of *individual* measurements are unpredictable on the quantum scale, measurements on a system in a given state yield a well-defined *distribution* of results. If we simply accept this notion and generalize our concept of "experiment" to the procedure of determining the distribution associated with many measurements of the same observable over a set of identically prepared systems, then we may maintain the idea that the *quantum state* of a system determines the results of all experiments measuring physical observables.

As we shall refer to the quantum state of a system often, we shall make a special symbol for it, $|\varphi\rangle$. The symbol $|\rangle$ is called a "ket." We may place different symbols inside the ket to distinguish different states. For instance, our system may be in the state $|\psi\rangle$ at time $t = 0$ and in the state $|\phi\rangle$ at some later time.

Keep in mind that, in general, $|\psi\rangle$ does not determine the result of any *individual* measurement of an observable, only probability distributions such as $\mathcal{P}(x)dx$, the probability of finding a particle with position x in the range $(x \to x + dx)$, or $\tilde{\mathcal{P}}(p)dp$, the probability of finding the momentum of the particle in the range $(p \to p + dp)$. In the case of photons we may also consider the observable of



spin or polarization and measure $P(\sigma)$, the probability of observing a right ($\sigma = 1$) or left ($\sigma = -1$) circularly polarized photon.

We mention the polarization of photons above to underscore the fact that we will keep the abstract portion of our discussion completely general. For the most part of the course, however, we will focus on systems of single particles which have no internal structure so that we will concern ourselves with only the observables of position $\vec{x}$, momentum $\vec{p}$ and those observables which may be derived from these two, such as energy $E = \frac{p^2}{2m} + V(\vec{x})$ and angular momentum $\vec{L} = \vec{x} \times \vec{p}$.

2 Specification of Quantum State

Classically, we know that knowledge of $\{x_i\}$ and $\{v_i\}$ at a given time specifies the state of the system. Quantum mechanically, the state of a system is specified when the probability distribution of all observables are determined. We still do not know what is the minimal amount of information needed to specify the quantum state $|\psi\rangle$. The answer to this question depends on the complexity of the system and the observables. As we mentioned at the end of the last section, we will focus on systems where all observables may be derived from the position and momentum of a single particle. For these systems we may ask whether we need specify both $\mathcal{P}(x)$ and $\tilde{\mathcal{P}}(p)$ to determine the quantum state. Perhaps, just one or the other will do. Or perhaps, some other quantity will specify both.

To address these issues consider the thought experiment carried out with photons in Figure 2.

Photons from a distance source S pass through a filter F, enter a semicircular chamber of radius R and eventually register on the photographic plate P. The thickness of the filter varies so that the intensity of photons entering the chamber at $y = 0$ is proportional to $F(x) = e^{\frac{-x^2}{2\sigma^2}}$. The geometry is arranged so that $R >> \sigma$. The amplitude of the entering wave $\psi(x)$ then obeys

$$\begin{aligned} |\psi(x)|^2 &\propto e^{\frac{-x^2}{2\sigma^2}} \\ \psi(x) &= \alpha e^{\frac{-x^2}{4\sigma^2}} e^{i\phi} \end{aligned} \tag{1}$$

We have included a proportionality constant and an arbitrary phase factor because (1) only determines $\psi(x)$ up to an arbitrary complex phase. In this situation we are fortunate and may determine the phase by assuming continuity of the phase of the wave as it crosses the filter so that $\phi = \frac{2\pi}{\lambda} \cdot x \sin\phi = (k \sin\phi)x \equiv k_x^s \phi$, where k_x^s is the x-component of the incoming momentum. Knowing the phase factor, we then have

$$\psi(x) = \alpha e^{\frac{-x^2}{4\sigma^2}} e^{ik_x^s x}.$$

Finally, if we wish to interpret $|\psi(x)|^2$ as the probability distribution $\mathcal{P}(x)$ for the location of individual phonons as they cross the filter, we may determine α from the normalization condition,

$$\begin{aligned} 1 &= \int \mathcal{P}(x)\, dx = \int |\psi(x)|^2\, dx = \int |\alpha|^2 e^{\frac{-x^2}{4\sigma^2}} |^2 |e^{ik_x^s x}|^2\, dx \\ &= \alpha^2 \int e^{\frac{-x^2}{2\sigma^2}}\, dx = \alpha^2 \sqrt{2\pi}\sigma \end{aligned}$$

thus,

$$\psi(x) = \frac{1}{\sqrt{\sqrt{2\pi}\sigma}} e^{-\frac{x^2}{4\sigma^2}} e^{ik_x^s x} \tag{2}$$

and

$$\mathcal{P}(x) = |\psi(x)|^2 = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \tag{3}$$

After entering the chamber with this wavefunction along x, the photons then propagate to the photographic plate where they arrive at the position of angle θ with an intensity given by

$$I \;\alpha\; \left| \int dx\, \psi(x) e^{-ixk \sin\theta} \right|^2 \tag{4}$$

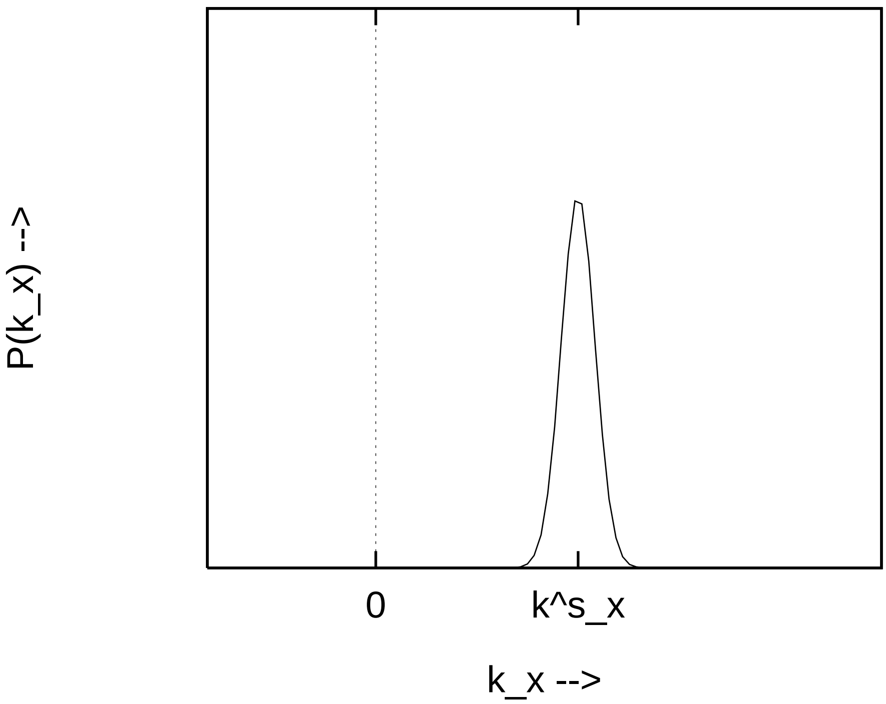
The phase factor $e^{-ixk\sin\theta}$ comes about because to arrive on the photographic plate at angle θ, photons leaving the plate at x must travel a distance $x \sin\theta$ *less* than those which cross at $x = 0$. (See figure.)

Note that the angle θ gives us a measurement of the x-component of the momentum of each photon as it left the filter to be recorded on the photographic plate. To arrive at angle θ the x component of the momentum must have obeyed $p_x = p \sin\theta$ where p was the total momentum given by $p = \frac{h}{\lambda} = \frac{h}{2\pi/k} = \hbar k$. Because the momentum p and wave number k are so closely related (just through the proportionality constant $\hbar$), we will often use them interchangeably. We for instance speak of the x component of $k, k_x \equiv p_x/\hbar = k \sin\theta$. If we now interpret the intensity (4) as a probability for individual photons we find

$$\tilde{\mathcal{P}}(k_x) \propto \left| \int dx\, \psi(x) e^{-ik_x x} \right|^2 \equiv |A|^2 \tag{5}$$

For our particular case,

$$\begin{aligned} A &\equiv \int dx\, \psi(x) e^{-ik_x x} \propto \int dx \left(e^{-\frac{x^2}{4\sigma^2}} e^{ik_x^s x} \right) e^{-ik_x x} \\ &= \int dx\, e^{-\frac{x^2}{4\sigma^2}} e^{-i(k_x - k_x^s)x} \end{aligned}$$



$$
\begin{aligned}
&= \int dx\, e^{-\frac{1}{4\sigma^2}x^2 - i\Delta x} \qquad ; \Delta \equiv k_x - k_x^s \\
&= \int dx\, e^{-\frac{1}{4\sigma^2}(x+2i\sigma^2\Delta)^2 - \sigma^2\Delta^2} \\
&= e^{-\sigma^2\Delta^2} \int dx\, e^{-\frac{1}{4\sigma^2}(x+2i\sigma^2\Delta)^2} \qquad ; u = x + 2i\sigma^2\Delta \\
&= e^{-\sigma^2\Delta^2} \int du\, e^{-\frac{1}{4\sigma^2}u^2} \\
&= e^{-\sigma^2\Delta^2} 2\sqrt{\pi}\sigma = e^{-\sigma^2(k_x-k_x^s)^2}\left(2\sqrt{\pi}\sigma\right) \qquad (6)
\end{aligned}
$$

So,

$$\tilde{\mathcal{P}}(k_x) \propto e^{-2\sigma^2(k_x-k_x^s)^2}$$

Normalizing, using the same procedure we used for $\mathcal{P}(x)$ gives the final result

$$\tilde{\mathcal{P}}(k_x) = \left(\sqrt{\frac{\pi}{2\sigma^2}}\right)^{-\frac{1}{2}} e^{-2\sigma^2(k_x-k_x^s)^2}$$

which appears as in the sketch below.

As a result in the localization in x, there is a spread in the momenta $\hbar k_x$ in accordance with the uncertainty principle. The *key lesson*, here, however comes from the fact that the distribution in the momentum in the x-direction is centered about its incoming value $\hbar k_x^s$, as we would expect on physical grounds. This is a direct result of the phase factor in (2). (The student should note that the inclusion of a *phase* in a Fourier integral of the form (6) always represents a *shift* in the final result). Without this phase factor it is impossible to determine where the distribution $\tilde{\mathcal{P}}(k_x)$ should be centered. Thus, $\mathcal{P}(x)$, which gives none of this phase information, is insufficient to specify the state of the system.

On the other hand, the function $\psi(x)$ specifies *both* $\mathcal{P}(x)$ and $\tilde{\mathcal{P}}(k_x)$ and so does specifies the state of the system entirely. We shall now see that the fact that $\psi(x)$ gives both $\mathcal{P}(x)$ and $\tilde{\mathcal{P}}(k_x)$ is a consequence of the general physical principle of *superposition.*

3 Principle of Superposition

We now return to our general abstract discussion. Before discussing the principle of superposition it is useful to establish the existence of a special class of states.

3.1 Pure States

A *pure state* of a system is always defined with respect to a physical observable. It is defined to be a state where measurement of the observable always yields the same result.

Such a state may be prepared by first measuring the observable and then maintaining only these systems in the experiment which yield a particular result. The state of the systems we keep, immediately after the sorting process and before any internal or external influences have a chance to have an effect, will be pure and always give back the same value if the observable is measured again.

3.2 Superposition

When a system is in a state where the measurement of an observable $\mathcal{O}$ may take on various values (as in the final location in a single state diffraction experiment) the state of the system somehow "contains," in a special quantum mechanical sense, the pure states corresponding to the alternative values of the observable. We say that the state of the system is a *superposition* of those pure states of the observable which represent the possible outcomes. As we can imagine a quantum state composed of any combination of pure states, the superposition of any two states leads to a new valid quantum state. We express the physical concept of the composition of an arbitrary state in terms of pure states symbolically as

$$|\psi\rangle = \sum_n |\mathcal{O}_n\rangle \tag{7}$$

where the $|\mathcal{O}_n\rangle$ represent pure states of the observable $\mathcal{O}_n$. Note that if the allowed values of $\mathcal{O}$ are continuous, then we would write an integral instead

$$|\psi\rangle = \int d\mathcal{O}\, |\mathcal{O}\rangle \tag{8}$$

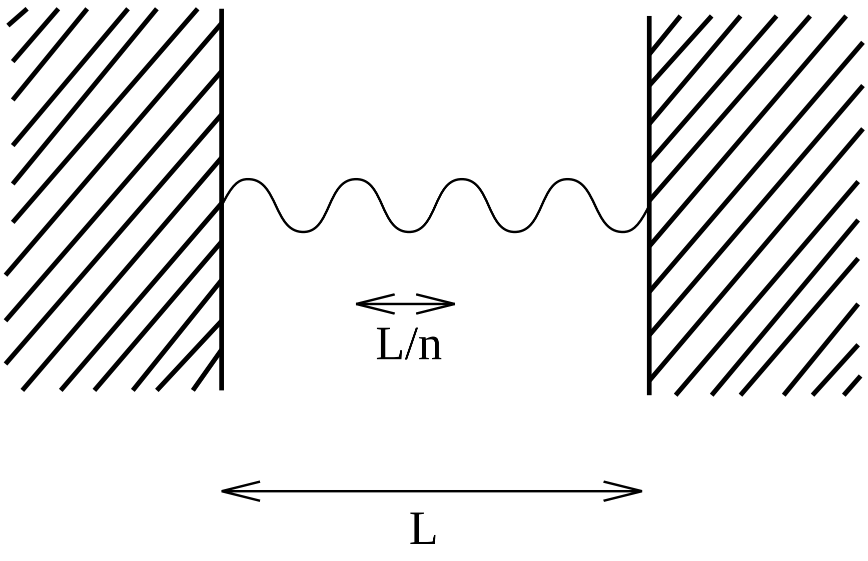
Because the measurement of an observable must always yield some value and a system in any state $|\psi\rangle$ may be measured, it is a general principle of quantum mechanics that all states may be written as superpositions of the pure states of any valid physical observable.

3.3 Simple Example of Superposition

Consider an apparatus as in Figure 3.3 consisting of two mirrors a distance L apart and containing a photon of wavelength $\lambda = \frac{L}{n}$. The wave function for this standing wave pattern looks something like

$$\phi(x) = A \sin \frac{2\pi}{\lambda} x = A \sin \frac{2\pi}{L} x = A \sin k\ x$$

where $k = \frac{3\pi}{L}$. If we are to measure the momentum of the photon (say by allowing it to collide with an electron), the de Broglie relations tells us that we will measure $p = \frac{h}{\lambda} = \hbar k$.



If we do the actual experiment, however, we will find something like in Figure 3.3

This is because we do not have a pure traveling wave, but a standing wave consisting of waves moving in two directions, $+x$ and $-x$.

We can see this mathematically by expanding $\sin k\ x$,

$$\phi(x) = A \sin k\ x = \frac{A}{2i} e^{ikx} - \frac{A}{2i} e^{-ikx}$$

The state of the photon is now the superposition of two pure states, one of momentum $\hbar k$ and one of momentum $-\hbar k$, each with weight equal in magnitude.

In our abstract language we would write the superposition as

$$|\phi\rangle = \frac{A}{2i} |\hbar k\rangle - \frac{A}{2i} |-\hbar k\rangle$$

so that when we measure p we will get one of the two pure states in the composition of $|\psi\rangle$ with probability somehow related to the weights in the composition. In this case, the weights are equal so we expect equal probabilities as observed in the experiment.

3.4 Fourier Transforms and Dirac δ-function

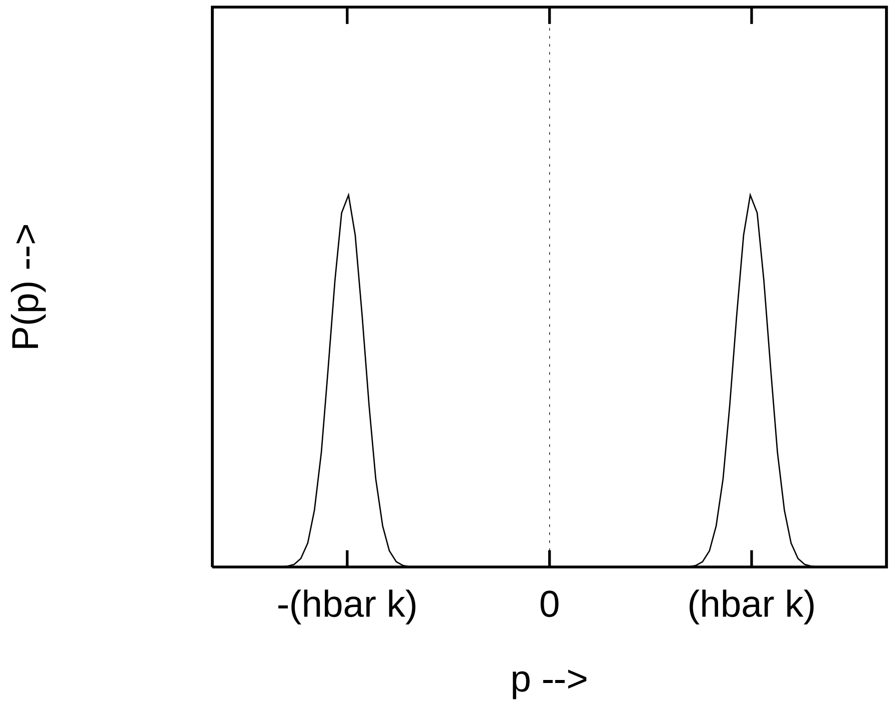
We will now use the physical principle of superposition, the *de Broglie* identification of e^{ikx} as a pure state of momentum $p = \hbar k$, and our filter scattering example to give a physical "proof" of the Fourier transform theorem.

The principle of superposition states that our state $|\psi\rangle$ in (2) is expressible as the superposition of pure momentum states. As momentum is continuous we write

$$|\psi\rangle = \int dk\, \tilde{\psi}(k)\ |\hbar k\rangle$$

where the $\tilde{\psi}(k)$ are the relative weights somehow related to the probability distribution $\tilde{p}(k)$ when measuring momentum p or equivalently $k = \frac{p}{\hbar}$. Writing this in our "concrete" language

$$\psi(x) = \int dk\, \tilde{\psi}(k) e^{ikx}. \tag{9}$$


P(p) -->
-(hbar k)
0
(hbar k)
p -->

Now, from (5) we had $\tilde{p}(k) = |A|^2$ where $A = \int dk\, \psi(x)e^{-ikx}$. One natural guess would be that

$$\tilde{\psi}(k) = \int dk\, \psi(x)e^{-ikx} \qquad (?)$$

We may verify this by inserting our guess into (9) and seeing if we recover $\psi(x)$.

$$\begin{aligned} I(x) &= \int dk \left[\int dx' \psi(x')e^{-ikx'}\right] e^{ikx} \\ &= \int dx' \psi(x') \int dk\, e^{ik(x-x')} = \int du \psi(x+u) \int dk\, e^{-iku} \end{aligned} \tag{10}$$

where we've made the change of variables $x' = x + u$.

3.4.1 Dirac Delta Function

We must now investigate the integral

$$2\pi\delta(u) = \int_{-\infty}^{\infty} dk\, e^{-iku} \tag{10.5}$$

To evaluate the integral we use a common trick in physics and insert a slight convergence factor $e^{-\epsilon|k|}$ which we later "turn off" by taking the limit $\epsilon \to 0$ at the end of the day. We then have,

$$\begin{aligned} 2\pi\delta(u) &= \int_{-\infty}^{\infty} dk\ e^{-\epsilon|k|}e^{-iku} \\ &= \int_{0}^{\infty} dk\ e^{-(\epsilon+iu)k} + \int_{-\infty}^{0} dk\ e^{(\epsilon-iu)k} \\ &= \frac{1}{\epsilon+iu} + \int_{0}^{\infty} dk\ e^{-(\epsilon-iu)k} \\ &= \frac{1}{\epsilon+iu} + \frac{1}{\epsilon-iu} = \frac{(\epsilon-iu)+(\epsilon+iu)}{(\epsilon+iu)(\epsilon-iu)} \\ &= \frac{2\epsilon}{\epsilon^2+u^2} \end{aligned}$$

So,

$$\delta(u) = \frac{1}{\pi}\frac{\epsilon}{\epsilon^2+u^2},$$

which appears like a very tall narrow spike of height $1/\pi\epsilon$ and width $\sim \epsilon$ as in Figure 3.4.1.

The height grows and width shrinks so that the area is always constant,

$$a = \int \delta(u)du = \int_{-\infty}^{\infty} \frac{1}{\pi}\frac{\epsilon}{\epsilon^2+u^2} d\epsilon = \frac{1}{\pi}\tan^{-1}\frac{u}{\epsilon}\Big|_{-\infty}^{\infty} = \frac{1}{\pi}\left(\frac{\pi}{2} - (-\frac{\pi}{2})\right) = 1$$

The δ-function has two basic properties from which others may be derived:

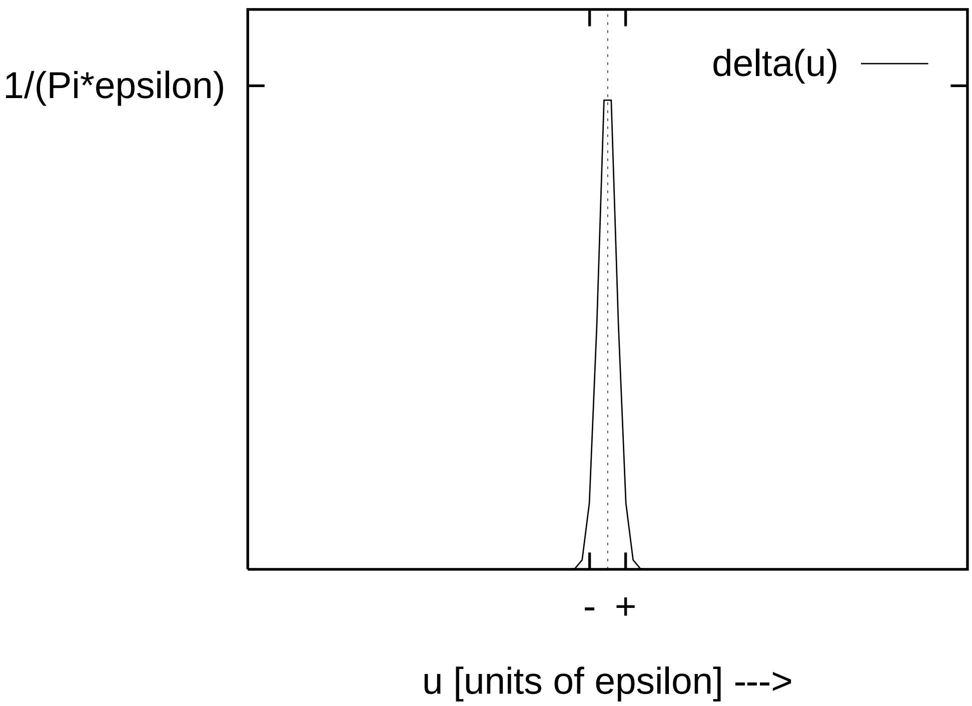
1. $\delta(u)$ vanishes outside of a tiny range ϵ.
2. The $\delta-$function has unit area. $\delta(u)du = 1$ (11)

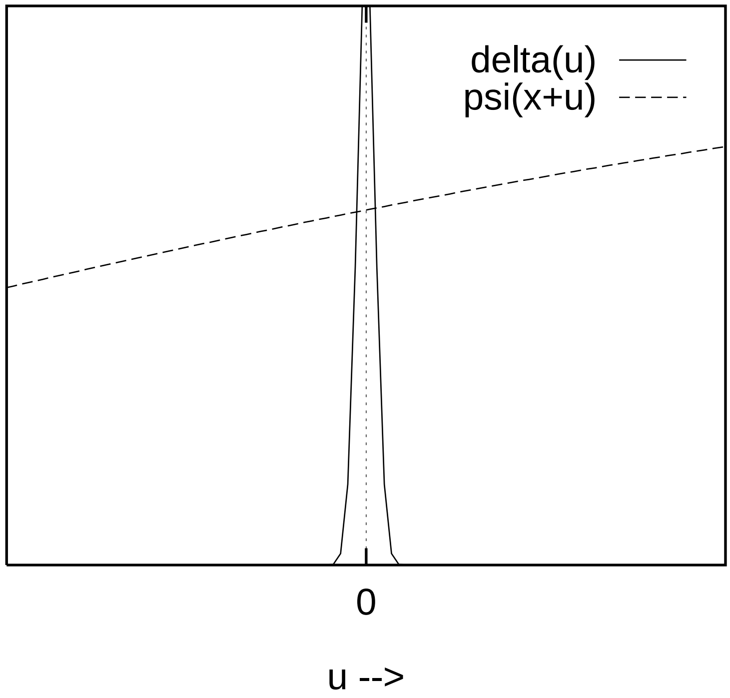
3.4.2 Fourier Transforms

Using the δ-function we may now complete the integral in (10),

$$I(x) = \int du\ \psi(x+u) \cdot 2\pi\delta(u) = 2\pi \int du\ \psi(x+u)\delta(u)$$

The integral appears as the product of the two functions in Figure 3.4.2.


1/(Pi*epsilon)
delta(u)
-
+
u [units of epsilon] --->


delta(u)
psi(x+u)
0
u -->

Because the peak may be made arbitrarily narrow, the values of u which contribute to the integral are so tiny that we may make the approximation $\psi(x+u) \simeq \psi(x)$ with a small error proportional to at most $\mathcal{O}(\epsilon)$ and which will vanish in the final limit $\epsilon \to 0$. Thus, we have

$$I(x) = 2\pi\psi(x) \int \delta(u) du = 2\pi\psi(x)$$

Our initial guess was nearly correct! We have found

$$2\pi\psi(x) = \int dk \left[\int dx \; e^{-ikx'} \psi(x') \right] e^{ikx}$$

or

$$\psi(x) = \int \frac{dk}{\sqrt{2\pi}} \left[\int \frac{dx'}{\sqrt{2\pi}} e^{-ikx} \psi(x') \right] e^{ikx}$$

This tells us that if we take the superposition

$$\psi(x) = \int \frac{dk}{\sqrt{2\pi}} \tilde{\psi}(k) e^{ikx} \equiv \mathcal{F}\left[\tilde{\psi}(k)\right] \tag{12}$$

we may find $\tilde{\psi}(k)$ from $\psi(x)$ from the inverse relationship

$$\tilde{\psi}(k) = \int \frac{dx}{\sqrt{2\pi}} \psi(x) e^{-ikx} \equiv \mathcal{F}^{-1}\left[\psi(x)\right] \tag{13}$$

(12) and (13) taken together are the Fourier inversion theorem, where the operation in (12), $\mathcal{F}\left[\tilde{p}si(k)\right]$ is known as the *Fourier transform*, while the operation in (13), $\mathcal{F}^{-1}\left[\psi(x)\right]$ is the *inverse Fourier transform*.

3.5 Interpretation of the Weights of Superposition

In (3.3) we were vague regarding the precise relationship between the weights in a superposition of pure states and the probability of observing each of the states in a measurement. Comparing (13) with (5), however, we see that to be consistent with the observations in our apparatus, the relationship is simply that:

The probability of obtaining a result when measuring an observable in a state composed of a superposition of pure states of the observable is proportional to the complex magnitude squared of the weight of the pure state observed.

Expressed in our concrete position-momentum example, this is

$$\tilde{p}(k) = |\tilde{\psi}(k)|^2 \tag{14}$$

3.6 Parseval's Theorem: Fourier Normalization

You will note that we wrote (14) as an equality rather than as a simple proportion. This reflects the fact that we have been very careful in our choice of constant factors in (12) and (13) so that if $\psi(x)$ is properly normalized,

$$1 = \int p(x) dx = \int |\psi(x)|^2 dx$$

So will be $\tilde{\psi}(k)$ when computed according to (13)

$$1 = \int \tilde{p}(k) dk = \int |\tilde{\psi}(k)|^2 dk.$$

We will now prove a slightly more general result known as Parseval's Theorem of which this is clearly a special case.

Parseval's Theorem:
$$\int \psi^*(x)\phi(x)dx = \int \tilde{\psi}^*(k)\tilde{\phi}(k)dk$$
To prove this, as with most theorems involving Fourier transforms, we need only use (10.5) and familiar integration techniques.

$$\begin{aligned}
\int e^{ikx}dk &= 2\pi\delta(x) \\
\int e^{ikx}dx &= 2\pi\delta(k)
\end{aligned}$$

We proceed as follows:

$$\begin{aligned}
\int \psi^*(x)\phi(x)dx &= \int \left[\int \frac{dk}{\sqrt{2\pi}} e^{ikx}\tilde{\psi}(k)\right]^* \left[\int \frac{dk'}{\sqrt{2\pi}} e^{ik'x}\tilde{\phi}(k')\right] dx \\
&= \int dk\tilde{\psi}^*(k) \int dk'\tilde{\phi}(k') \int \frac{dx}{2\pi} e^{ix(k'-k)} \\
&= \int dk\tilde{\psi}^*(k) \int dk'\phi(k')\delta(k'-k) \\
&= \int dk\tilde{\psi}^*(k) \int dq\tilde{\phi}(k+q)\delta(q); \quad q = k'-k \\
&= \int dk\tilde{\psi}^*(k)\tilde{\phi}(k) \int dq\delta(q) \\
&= \int dk\tilde{\psi}^*(k)\tilde{\phi}(k)
\end{aligned}$$

3.7 Representations of Quantum States/Hermitian Inner Product

As we have seen, $\psi(x)$ is sufficient to compute the distributions for position $p(x)$ and momentum $\tilde{p}(p)$ through (3), (13) and (14) and thus is sufficient to specify or represent the quantum state $|\psi\rangle$ of a system where x and p are the only basic observables. However, through (12) and (13) we see that $\psi(x)$ and $\tilde{\psi}(k)$ provide equivalent information and so either is an equally valid way of representing the state $|\psi\rangle$. If we wish, we can, in principle, compute all of our physical quantities using $\psi(x)$ in the *position representation*, or using $\tilde{\psi}(k)$ in the *momentum representation* and get equally valid results.

We have already seen an example of this in Parseval's Theorem. We may compute the *Hermitian Inner product* between two states $|\psi\rangle$ and $|\phi\rangle$, either in the position or momentum representation and get the same result
$$I = \int \psi^*(x)\phi(x)dx = \int \tilde{\psi}^*(k)\tilde{\phi}(k)dk$$
Because this quantity is the same regardless of which representation we choose, it depends only on the states $|\psi\rangle$ and $|\phi\rangle$ and we thus give it a special symbol independent of the choice of representation to remind us of this fact
$$\langle\psi|\phi\rangle \equiv I = \int \psi^*(x)\phi(x)dx = \int \tilde{\psi}^*(k)\tilde{\phi}(k)dk \tag{15}$$
When we write $\langle\psi|\phi\rangle$ it is clear what is meant. To evaluate it, the reader is responsible for picking a convenient representation of his own choosing and then performing the appropriate integration in (15).

The Hermitian Inner product integral itself has many interesting mathematical properties and so it is useful to have a notation for it as well. We shall write
$$\int \psi^*(x)\phi(x)dx \equiv (\psi, \phi) \tag{16}$$

and

$$\int \tilde{\psi}^*(k)\tilde{\phi}(k)dk \equiv (\tilde{\psi},\tilde{\phi})$$

So that (f,g) just means form the product f^*g of the two functions and integrate over their argument. (15) then is written compactly as

$$\langle\psi|\phi\rangle \equiv (\psi,\phi) = (\tilde{\psi},\tilde{\phi})$$

We note that the mathematical structure of the Hermitian Inner product (16) as a sum over the product of values taken from two "lists" is very similar to that of the dot product between two vectors $\vec{A}\cdot\vec{B} = \sum_i A_i \cdot B_i$. It is therefore not surprising that they share many properties. The most important of these (whose proof we defer to problem set 6) is the Cauchy-Schwartz inequality. For vectors we have

$$\begin{aligned}
(\vec{A}\cdot\vec{B})^2 &= \left(||\vec{A}||||\vec{B}||\cos\theta\right)^2 = ||\vec{A}||^2||\vec{B}||^2\cos^2\theta \le ||\vec{A}||^2||B||^2 = \left(\vec{A}\cdot\vec{A}\right)\left(\vec{B}\cdot\vec{B}\right) \\
\left(\vec{A}\cdot\vec{B}\right)^2 &\le \left(\vec{A}\cdot\vec{A}\right)\left(\vec{B}\cdot\vec{B}\right)
\end{aligned}$$

For the Hermitian Inner product, you will prove

$$|(\psi,\phi)|^2 \le (\psi,\psi)(\phi,\phi).$$