A17.2 The Uncertainty Principle

The uncertainty principle has a technical statement and derivation along with an *interpretation* or explanation of what the technical statement implies. To illustrate, suppose we have either a light-like wave function or an electron-like wave function that goes through a narrow slit. The particle-like wave function is initially traveling in the z direction and the slit is in the y direction. After the wave goes through the slit, it spreads out in the *x*-direction. Because of this, if one measures the *x* position far from the slit many times, one will get a spread in values, even if the measurements are exact. This is just the nature of waves. The same conclusion holds for momentum; if one measures the momentum in the *x* direction, one will get a spread in values, even if the measurements are exact.

Technical statement of the uncertainty relation.

It is remarkable that by the use of standard vector-space arguments (with no reference to particles), one can get a relation between the spreads in position and momentum. It is (see below for a proof):

$$\langle (p_x - \langle p_x \rangle)^2 \rangle \, \langle (\mathrm{x} - \langle x \rangle)^2 \rangle \geq \hbar^2/4 \tag{1}$$

where the $\hbar$ comes from the fact that $[p_x, x] = -i\hbar$. The symbol $\langle O \rangle$ signifies the expectation value or average value of whatever is inside them. This average corresponds to the quantum mechanical average value, $\langle \psi | O | \psi \rangle$, one would obtain if one measured the quantity corresponding to *O* many times in the state $|\psi\rangle$. If we take the square root of this relation and use root-mean-square quantities, we get the usual form of the uncertainty principle

$$(\Delta p_x)_{rms} (\Delta \mathrm{x})_{rms} \geq \frac{\hbar}{2} \tag{2}$$

Usual interpretation of the uncertainty relation.

This relation is often interpreted as follows: *There exists a particle embedded in the wave function whose position and momentum cannot be simultaneously measured exactly. The spread in measuring momentum times the spread in measuring position is greater than or equal to* $\hbar/2$*.* This is quite mysterious if one thinks of actual particles. Why should we not be able to simultaneously measure position and momentum to arbitrary accuracy?

The answer is that the underlying assumption—that there are actual particles (with definite position and momentum)—is not warranted because there is no evidence for particles. The only thing the uncertainty principle tells us is that if we make many measurements of position and then, on separate runs of the experiment, we make many measurements of momentum, the product of the rms spreads of the two quantities must be greater than $\hbar/2$. This result, as we show below, is simply a *mathematical consequence* of the (vector) nature of wave functions. Thus, if there are no particles, the uncertainty principle loses all its mystery because it is simply a mathematical property of waves (wave functions).

> [Digression: I believe this usual interpretation is incorrect, and not just because of the use of the concept of particles. In Eq. (1), we see that it is the expectation value of $(\delta p_x)^2$ times the expectation value of $(\delta x)^2$. That means one does many measurements on $(\delta p_x)^2$ and then, **separately**, one does many experiments on $(\delta x)^2$, and then one multiplies them together. It does *not* say to do many *simultaneous* measurements of $(\delta p_x)^2$ and $(\delta x)^2$. That would correspond to $\langle (p_x - \langle p_x \rangle)^2 (x - \langle x \rangle)^2 \rangle$.]
>
> Note: The uncertainty principle is often used is to say that if the wave function is confined to a very small region in space, it will have a large spread in momentum. This is a valid translation of the import of Eqs. 1 and 2. It is saying that if the Fourier transform of the wave function with respect to position is sharply peaked, then the Fourier transform of the wave function with respect to momentum will be spread out.

Proof of the uncertainty principle.

Vectors. Scalar product.

We start with a review of vector mathematics. Suppose $\phi(q)$ and $\psi(q)$ are functions of q. Then their scalar product is defined as

$$\langle \phi | \psi \rangle = \int dq \;\; \bar{\phi}(q) \psi(q) \tag{3}$$

where the bar denotes the complex conjugate. If $\psi(q)$ is a column vector, say $\Psi = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$, then the adjoint vector is a row vector

$$\Psi^\dagger = [\bar{\psi}_1, \bar{\psi}_2] \tag{4}$$

In that case the scalar product is

$$\langle \phi | \psi \rangle = \int dq \, \phi^\dagger(q) \psi(q) = \int dq \, \bar{\phi}_i(q) \psi_i(q) \tag{5}$$

where a sum over the repeated index is implied. The expectation value for an operator O when the system is in state ψ is

$$\langle \psi | O | \psi \rangle = \langle \mathrm{O} \rangle = \int dq \, \psi^\dagger(q) O \psi(q) \tag{6}$$

Adjoint of an operator.

If ψ is a column vector and O an operator, then the adjoint operator, $O^\dagger$ is defined by

$$[O\psi]^\dagger = \psi^\dagger O^\dagger \tag{7}$$

one can see from this definition and the complex conjugation in Eq. (4), that if O is a matrix, then

$$[O^\dagger]_{ij} = \bar{O}_{ji} \tag{8}$$

A somewhat more general way than Eq. (7) to define the adjoint is to suppose that for any two states ϕ and ψ, the adjoint operator is the operator that obeys

$$\int dq\ \ \phi^\dagger(O\psi) = \int dq(O^\dagger\phi)^\dagger\psi \tag{9}$$

From this, using integration by parts (under the appropriate conditions) one deduces that the adjoint of the derivative is minus the derivative,

$$\left(\frac{\partial}{\partial x}\right)^\dagger = -\frac{\partial}{\partial x} \tag{10}$$

We note that it follows from these definitions that

$$(A^\dagger)^\dagger = A \tag{11}$$

The proof. (From Max Born, Atomic Physics, 6th ed., 1957.)
First we note that for any operator *A*,

$$\begin{aligned} \langle AA^\dagger\rangle &= \int dq\ \ \psi^\dagger(AA^\dagger\psi) \\ &= \int dq\ \ (A^\dagger\psi)^\dagger(A^\dagger\psi) \\ &= \int dq\,|A^\dagger\psi|^2 \geq 0 \end{aligned} \tag{12}$$

Now let λ be any real number and *A*, *B* be hermitian operators so that $A^\dagger = A, B^\dagger = B$. Then using $(A + iB)^\dagger = A^\dagger - iB^\dagger$, we have

$$J(\lambda) = \langle (A + i\lambda B)^\dagger(A + i\lambda B)\rangle = \langle A^2\rangle + \lambda^2\langle B^2\rangle - i\lambda\langle AB - BA\rangle \geq 0 \tag{13}$$

One can minimize this with respect to λ and find that $J(\lambda)$ at the min is

$$J(\lambda)_{min} = \langle A^2\rangle + \frac{\langle AB - BA\rangle^2}{4\langle B^2\rangle} \geq 0 \tag{14}$$

We now set

$$A = p_x - \langle p_x\rangle = -i\,\hbar\frac{\partial}{\partial x} - \langle -i\,\hbar\frac{\partial}{\partial x}\rangle \tag{15}$$

$$B = x - \langle x \rangle$$

so that

$$[AB - BA] = -i\hbar \tag{16}$$

Then the inequality of Eq. (14), with a little algebraic manipulation (and with the wave function normalized of course), becomes the uncertainty relation of Eq. (1).