Chapter 11. Commutators and Measurements

Section 11.4. A General Theorem About Uncertainty in Measurements

Section 11.1. Observables Whose Operators Commute

§ 1 *Introduction.* A pure state of a hydrogen atom is described, in those cases in which we *can ignore spin*, by a ket $|n,\ell,m\rangle$. If the system is in this state and we measure the energy, the result is sure to be (see Metiu, *Quantum Mechanics*)

$$E_n = -\frac{1}{n^2}\,\frac{\mu e^4}{2(4\pi\epsilon_0)^2\hbar^2} \tag{1}$$

If we measure the square of the angular momentum (which is proportional to the rotational energy of the electron), we are sure to obtain the value

$$\hbar^2\ell(\ell+1) \tag{2}$$

If we measure the projection of the angular momentum on the OZ axis, we are sure to obtain

$$\hbar m \tag{3}$$

This happens because $|n, \ell, m\rangle$ is an eigenstate (hence a pure state) of three operators, the energy $\hat{H}$, the square of the angular momentum $\hat{L}^2$, and the projection $\hat{L}_z$ of the angular momentum on the OZ axis. In other words,

$$\hat{H}|n, \ell, m\rangle = E_n|n, \ell, m\rangle \tag{4}$$

$$\hat{L}^2|n, \ell, m\rangle = \hbar^2\ell(\ell+1)|n, \ell, m\rangle \tag{5}$$

$$\hat{L}_z|n, \ell, m\rangle = \hbar m|n, \ell, m\rangle \tag{6}$$

If the system is in the state $|n, \ell, m\rangle$ and we want to determine the position of the electron, we can only know the probability

$$P_{n,\ell,m}(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = |\langle x, y, z \,|\, n, \ell, m\rangle|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{7}$$

that the electron is inside the cube of volume $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ centered on the point $\{x, y, z\}$. The exact position is uncertain! Why is position different from $\hat{H}$, $\hat{L}^2$, or $\hat{L}_z$? Why can we be certain of energy and the magnitudes of L^2 and L_z but cannot be certain of the position of the particle?

When you first learned about the pure energy states of the hydrogen atom, you were probably told that $\hat{H}$, $\hat{L}^2$, and $\hat{L}_z$ commute with each other and that this is why they have common eigenfunctions. We can go a step further and suggest that $|n, \ell, m\rangle$ is not an eigenstate of the position operator or of the momentum operator because those operators *do not commute* with $\hat{H}$, $\hat{L}^2$, and $\hat{L}_z$. There seems to be something special about Hermitian operators that commute with each other. The following theorem tells us what that special quality is.

§ 2 *Commuting operators and their eigenstates.* In what follows, we discuss and prove the following theorem.

Consider two operators $\hat{A}$ and $\hat{B}$, defined in the same space and representing observables.

(a) If $\hat{A}$ and $\hat{B}$ have the same eigenstates, then $\hat{A}$ and $\hat{B}$ commute.

(b) If $\hat{A}$ and $\hat{B}$ commute then:

(b1) All *nondegenerate* eigenstates of $\hat{A}$ are also eigenstates of $\hat{B}$ and vice versa.

(b2) For each *degenerate* eigenvalue α of $\hat{A}$, of multiplicity n, we can construct kets $|\phi_m\rangle$, $m = 1, \ldots, n$, that are eigenstates of $\hat{A}$ corresponding to the eigenvalue α and are also eigenstates of $\hat{B}$. This means that

$$\hat{A}|\phi_m\rangle = \alpha|\phi_m\rangle \text{ for } m = 1, 2, \ldots, n \tag{8}$$

and

$$\hat{B}|\phi_m\rangle = \lambda_m|\phi_m\rangle \text{ for } m = 1, 2, \ldots, n \tag{9}$$

In other words, commuting operators do have common eigenstates, but we need to do some extra work to find them if degeneracy is present. The theorem is silent about the eigenvalues λ_m in (b2).

We will spend quite a bit of time understanding the meaning and the consequences of this theorem (§3 and §4) and then prove it in §5.

§ 3 *An example: the harmonic oscillator and the reflection operator.* Consider a harmonic oscillator with the Hamiltonian

$$\hat{H} = a\frac{\mathrm{d}^2}{\mathrm{d}x^2} + bx^2 \tag{10}$$

where a and b are two constants whose values are of no interest here.

We define the operator $\hat{\sigma}$ through

$$\hat{\sigma}\psi(x) \equiv \psi(-x) \tag{11}$$

We can see by direct calculation that $\hat{\sigma}$ commutes with $\hat{H}$:

$$\begin{aligned} \hat{H}\hat{\sigma}\psi(x) &= a\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(-x) + bx^2\psi(-x) & (12)\\ \hat{\sigma}\hat{H}\psi(x) &= \hat{\sigma}\left(a\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) + bx^2\psi(x)\right) \\ &= a\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(-x) + bx^2\psi(-x) & (13) \end{aligned}$$

We used the facts that $\mathrm{d}^2/\,\mathrm{d}x^2$ and x^2 are not changed by $\hat{\sigma}$. Eqs. 12 and 13 tell us that $\hat{H}$ and $\hat{\sigma}$ commute when acting on $\psi(x)$. Since this is valid for any function $\psi(x)$, we conclude that

$$\hat{H}\hat{\sigma} = \hat{\sigma}\hat{H} \tag{14}$$

The two operators commute and we can examine whether they satisfy the statements made in the theorem given in §2.

It is easy to find the eigenvalues and the eigenvectors of $\hat{\sigma}$. We have

$$\hat{\sigma}^2\psi(x) = \hat{\sigma}\hat{\sigma}\psi(x) = \hat{\sigma}\psi(-x) = \psi(x)$$

This equation tells us that the eigenvalue of $\hat{\sigma}^2$ is 1. As a result, $\hat{\sigma}$ has the eigenvalues $+1$ and -1. Because the eigenvalues are real numbers, $\hat{\sigma}$ is

a Hermitian operator. One can easily show that $+1$ and -1 are the only eigenvalues.

What are the eigenfunctions of $\hat{\sigma}$? *Any* function f having the property

$$f(-x) = -f(x) \tag{15}$$

(an anti-symmetric function) satisfies the eigenvalue equation

$$\hat{\sigma} f(x) = f(-x) = -f(x) \tag{16}$$

Thus we see that if f is antisymmetric, it is an eigenvector of $\hat{\sigma}$ corresponding to the eigenvalue -1.

Any function g with the property

$$g(-x) = g(x) \tag{17}$$

(a symmetric function) is an eigenstate of $\hat{\sigma}$ with the eigenvalue $+1$, since

$$\hat{\sigma} g(x) = g(-x) = g(x) \tag{18}$$

Exercise 1 Which of xe^{-x^2}, $x^2e^{-x^2}$, xe^{-x}, ..., $x^{2n+1}e^{-x^2}$, $x^{2n}e^{-x^2}$, ... are eigenstates of $\hat{\sigma}$ with the eigenvalue -1 and which are eigenstates with the eigenvalue 1?

All symmetric functions are eigenstates of $\hat{\sigma}$ with eigenvalue $+1$. All anti-symmetric functions are eigenstates of $\hat{\sigma}$ with eigenvalue -1.

Let us examine how these observations relate to our theorem. You can look up the eigenstates $\phi_n(x)$ (where $x = r - r_0$ and r_0 is the equilibrium

position of the oscillator) of the harmonic oscillator (see Metiu, *Quantum Mechanics*, page 253) and verify the following property:

$$\begin{aligned} \hat{\sigma}\phi_n(x) \equiv \phi_n(-x) &= -\phi_n(x) \quad \text{if } n \text{ is odd} \\ \hat{\sigma}\phi_n(x) \equiv \phi_n(-x) &= \phi_n(x) \quad \text{if } n \text{ is even} \end{aligned} \tag{19}$$

Every eigenstate of $\hat{H}$ is an eigenstate of $\hat{\sigma}$. This is in agreement with our theorem, which states that because the eigenvalues of $\hat{H}$ are not degenerate and $[\hat{H}, \hat{\sigma}] = 0$, all eigenstates of $\hat{H}$ must be eigenstates of $\hat{\sigma}$.

Now let us check whether the eigenfunctions of $\hat{\sigma}$ are eigenfunctions of $\hat{H}$. For example, x^2, x^4, x^6, ... are eigenfunctions of $\hat{\sigma}$ corresponding to the eigenvalue $+1$. They are definitely *not* eigenfunctions of $\hat{H}$. What went wrong? It is here that the phrase "defined in the same space" comes into play. The wave functions $\phi_n(x)$ of the harmonic oscillator must satisfy

$$\int_{-\infty}^{+\infty} \phi_n(x)^* \, \phi_n(x) \, \mathrm{d}x = 1; \tag{20}$$

they belong to the space L^2. The functions x^2, x^4, x^6, ... do not satisfy Eq. 20; they don't belong to L^2 and are of no interest to us.

The eigenstates $\phi_n(x)$ of $\hat{H}$ form a complete basis set and the space in which $\hat{H}$ acts is the space of all functions that can be written in the form

$$\psi(x) = \sum_{n=0}^{\infty} c_n \phi_n(x) \tag{21}$$

where the coefficients c_n are complex numbers. The theorem requires that $\hat{\sigma}$ must be defined in this space.

Because $\phi_n(x)$ is symmetric when n is an even (nonnegative) integer, all symmetric functions in this space are given by

$$s(x) \equiv \sum_{k=0}^{\infty} c_k \phi_{2k}(x) \tag{22}$$

where $c_0, c_1, \ldots$ are arbitrary complex numbers. The anti-symmetric functions are given by

$$a(x) \equiv \sum_{k=0}^{\infty} b_k \phi_{2k+1}(x) \tag{23}$$

where $b_0, b_1, \ldots$ are arbitrary complex numbers. Functions such as x^2 and x^4, that gave us trouble, do not belong to this space (they are not normalizable).

Exercise 2 I claim that the b_k coefficients in Eq. 23 must satisfy $\sum_{k=0}^{\infty} b_k^* b_k < \infty$. Explain why that claim is true.

Obviously

$$\hat{\sigma} s(x) = s(x)$$

and

$$\hat{\sigma} a(x) = -a(x)$$

so $s(x)$ and $a(x)$ are eigenfunctions of $\hat{\sigma}$. Nevertheless, they are *not* eigenfunctions of $\hat{H}$. It seems that the theorem is wrong! However, it is not. It does not say that *any* eigenstate of $\hat{\sigma}$ must be an eigenstate of $\hat{H}$. It is more modest: it says that it is *possible to find a complete set of orthonormal states* that are pure states of both operators. In this example, these states are $\phi_n(x)$, $n = 0, 1, 2, \ldots$. They are eigenstates of both operators and they form a complete orthonormal basis set for the space on which both operators act.

§ 4 *The states of the hydrogen atom.* The theorem works fine when we examine the eigenstates $|n, \ell, m\rangle$ of the hydrogen atom. The eigenstates

$|n, \ell, m\rangle$ are common eigenstates of the commuting operators $\hat{H}$, $\hat{L}^2$, and $\hat{L}_z$. Every eigenstate of $\hat{H}$ is an eigenstate of $\hat{L}^2$ and an eigenstate of $\hat{L}_z$. But this happens *because we constructed* that states $|n, \ell, m\rangle$ to have these properties. Consider the state

$$|\psi_{n,\ell}\rangle = \sum_{m=-\ell}^{+\ell} c_m |n, \ell, m\rangle \tag{24}$$

where the c_m are arbitrary complex numbers. You can easily show that this is an eigenstate of $\hat{H}$ and of $\hat{L}^2$. However, *it is not an eigenstate of* $\hat{L}_z$, because

$$\hat{L}_z|\psi_{n,\ell}\rangle = \sum_{m=-\ell}^{+\ell} c_m \hbar m |n, \ell, m\rangle \neq \hbar m \sum_{m=-\ell}^{+\ell} c_m |n, \ell, m\rangle \tag{25}$$

We see that not all eigenstates of $\hat{H}$ and $\hat{L}^2$ are eigenstates of $\hat{L}_z$, even though all three operators commute. However, we can find a set of eigenfunctions (namely, $|n, \ell, m\rangle$) that are common to all three operators, as the theorem specifies.

Exercise 3 Show that $|\psi\rangle$ defined by Eq. 24 is an eigenstate of $\hat{H}$ and of $\hat{L}^2$.

Exercise 4 For the hydrogen atom, are the states 2s, $2p_x$, $2p_y$, and $2p_z$ eigenstates of $\hat{H}$, $\hat{L}^2$, and $\hat{L}_z$? The definitions of 2s, $2p_x$, $2p_y$, and $2p_z$ are given in any introductory book on quantum mechanics (see Metiu, page 321).

---

§ 5 *The proof of the theorem.* Now that we have clarified what the theorem says, let us prove it.

It is easy to show that two operators $\hat{A}$ and $\hat{B}$ that have the same eigenstates must commute. Let $|u_i\rangle$, $i = 1, 2, \ldots$, be the eigenstates of both operators, with corresponding eigenvalues α_i for $\hat{A}$ and β_i for $\hat{B}$Therefore the operators are

$$\hat{A} = \sum_{i=1}^{\infty} |u_i\rangle \, \alpha_i \, \langle u_i| \tag{26}$$

and

$$\hat{B} = \sum_{i=1}^{\infty} |u_i\rangle \, \beta_i \, \langle u_i| \tag{27}$$

To show that $\hat{A}\hat{B} - \hat{B}\hat{A} = 0$, we calculate the two products. First,

$$\begin{aligned} \hat{A}\hat{B} &= \left(\sum_{i=1}^{\infty} |u_i\rangle \, \alpha_i \, \langle u_i| \right) \left(\sum_{j=1}^{\infty} |u_j\rangle \, \beta_j \, \langle u_j| \right) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |u_i\rangle \, \alpha_i \, \langle u_i \,|\, u_j\rangle \, \beta_j \, \langle u_j| \end{aligned} \tag{28}$$

Since $\langle u_i \,|\, u_j\rangle = \delta_{ij}$, it follows that

$$\hat{A}\hat{B} = \sum_{i=1}^{\infty} |u_i\rangle \, \alpha_i \beta_i \, \langle u_i| \tag{29}$$

A similar calculation shows that $\hat{B}\hat{A} = \sum_{i=1} |u_i\rangle \, \beta_i \alpha_i \, \langle u_i| = \sum_{i=1} |u_i\rangle \, \alpha_i \beta_i \, \langle u_i|$. We have found that if $\hat{A}$ and $\hat{B}$ have the same eigenstates then $\hat{A}\hat{B} = \hat{B}\hat{A}$, which is what we wanted to prove.

Now let us assume that $[\hat{A}, \hat{B}] = 0$.

Suppose that $|u\rangle$ is an eigenstate of $\hat{A}$, corresponding to eigenvalue α:

$$\hat{A}|u\rangle = \alpha|u\rangle \tag{30}$$

Because $\hat{A}\hat{B} = \hat{B}\hat{A}$, we have

$$\hat{A}\hat{B}|u\rangle = \hat{B}\hat{A}|u\rangle = \alpha\hat{B}|u\rangle \tag{31}$$

This means that $\hat{B}|u\rangle$ is an eigenstate of $\hat{A}$ corresponding to the eigenvalue α.

If this eigenvalue *is not degenerate*, $\hat{B}|u\rangle$ must be proportional to $|u\rangle$. Denoting the proportionality constant by β, we can write

$$\hat{B}|u\rangle = \beta|u\rangle \tag{32}$$

This means that the eigenstate $|u\rangle$ of $\hat{A}$ is also an eigenstate of $\hat{B}$. Reversing the roles of $\hat{A}$ and $\hat{B}$, we can conclude similarly that when $[\hat{A}, \hat{B}] = 0$, all *nondegenerate* eigenstates of $\hat{B}$ are eigenstates of $\hat{A}$.

If the eigenstate $|u\rangle$ of $\hat{A}$ *corresponds to a degenerate eigenvalue* α, this argument falls apart and we have to dig deeper. We have already shown that $\hat{B}|u\rangle$ is an eigenstate of $\hat{A}$ corresponding to α, but the argument leading to Eq. 32 no longer works.

Let n be the multiplicity of the eigenvalue α and denote by $|u_1\rangle$, $|u_2\rangle$, $\ldots$, $|u_n\rangle$ the set of *linearly independent* eigenstates of $\hat{A}$ corresponding to α. Any linear combination of these eigenstates,

$$|\psi\rangle = \sum_{j=1}^{n_i} c_j \, |u_j\rangle \tag{33}$$

is an eigenstate of $\hat{A}$ corresponding to the eigenvalue α, as shown below.

$$\hat{A}|\psi\rangle = \sum_{j=1}^{n} c_j \, \hat{A}|u_j\rangle = \alpha \sum_{j=1}^{n} c_j \, |u_j\rangle = \alpha|\psi\rangle \tag{34}$$

Conversely, any eigenstate of $\hat{A}$ corresponding to α can be written in the form in Eq. 33.

The set of *all eigenstates* of $\hat{A}$ defines a linear space that I will denote by $\mathcal{S}$. The set of all vectors given by an expression like Eq. 33 forms a subspace $\mathcal{S}(\alpha)$. All elements of $\mathcal{S}(\alpha)$ belong to $\mathcal{S}$. As we have just seen, every element of $\mathcal{S}(\alpha)$ is an eigenstate of $\hat{A}$ corresponding to the eigenvalue α. Eq. 34 tells us that when $\hat{A}$ acts on an element of $\mathcal{S}(\alpha)$, it produces an element of $\mathcal{S}(\alpha)$.

But we have shown that $\hat{B}|u_i\rangle$ satisfies $\hat{A}(\hat{B}|u_i\rangle) = \alpha(\hat{B}|u_i\rangle)$ and therefore $\hat{B}|u_i\rangle$ is an element of $\mathcal{S}(\alpha)$. This means that there are complex numbers $\gamma_{i,1}, \ldots, \gamma_{i,n}$ for which

$$\hat{B}|u_i\rangle = \sum_{k=1}^{n} \gamma_{i,k}|u_k\rangle \tag{35}$$

is valid. It follows that when $\hat{B}$ acts on a vector in the subspace $\mathcal{S}(\alpha)$, it produces a vector in the same subspace. In other words, because $\hat{A}$ and $\hat{B}$ commute, the subspace $\mathcal{S}(\alpha)$ is invariant to both of them.

Consider now the matrix B with elements

$$\langle u_j \,|\, \hat{B} \,|\, u_i\rangle, \quad i, j = 1, \ldots, n_i$$

We can solve the eigenvalue problem for B to find its eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors, and then use the eigenvectors to calculate the kets $|\phi_1\rangle$, $\ldots$, $|\phi_n\rangle$ that are eigenstates of $\hat{B}$,

$$\hat{B}|\phi_m\rangle = \lambda_m|\phi_m\rangle, \quad m = 1, \ldots, n$$

The kets $|\phi_1\rangle, \ldots, |\phi_n\rangle$ belong to $\mathcal{S}(\alpha)$ by construction, so we also have

$$\hat{A}|\phi_m\rangle = \alpha|\phi_m\rangle, \quad m = 1, \ldots, n$$

This proves the theorem and tells us how to construct eigenstates of $\hat{B}$ that are also eigenstates of $\hat{A}$, when the eigenstates of $\hat{A}$ are degenerate. The

degenerate eigenstates of $\hat{A}$ are not automatically eigenstates of $\hat{B}$, but we can always construct kets that are eigenstates of both operators.

§ 6 *A simple example.* Let us test the theorem for the two operators represented by the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{36}$$

and

$$B = \begin{pmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} \tag{37}$$

Obviously they are Hermitian. To save mental effort, I have done all the required calculations in the `Mathematica` file `WorkBook11.nb`.

To see if the matrices commute, I calculate

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} \tag{38}$$

and

$$BA = \begin{pmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} \tag{39}$$

We have $AB = BA$. We did not actually need to do this calculation since A is the unit matrix (representing the unit operator), which commutes with all matrices.

The eigenvectors of A are (see `WorkBook 11`)

$$\left.\begin{aligned} v(1) = \{0, 1\} \quad \text{with eigenvalue } a_1 = 1 \\ v(2) = \{1, 0\} \quad \text{with eigenvalue } a_2 = 1 \end{aligned}\right\} \tag{40}$$

They are degenerate. The components of $v(j)$ are denoted by $v_i(j)$, $i = 1, 2$, so that, for example, $v_1(1) = 0$ and $v_2(1) = 1$.

The eigenvectors and eigenvalues of B are (see `WorkBook 11`)

$$\left.\begin{array}{ll} w(1) = \{\mathrm{i}, 1\} & \text{with eigenvalue } b_1 = 2 \\ w(2) = \{-\mathrm{i}, 1\} & \text{with eigenvalue } b_2 = 0 \end{array}\right\} \tag{41}$$

The eigenvectors $w(1)$ and $w(2)$ of B are not degenerate.

Part (b1) of the theorem says that, because the eigenstates of B are not degenerate, they must be eigenstates of A. I verified in `WorkBook 11` that this statement is correct, namely $Aw(1) = w(1)$ and $Aw(2) = w(2)$.[1]

Next I tested whether the degenerate eigenvectors of A are eigenvectors of B. They are not (see WorkBook 11). The first statement in part (b2) of the theorem is correct.

Part (b2) also tells us that while it is not compulsory for the degenerate eigenstates of A to be eigenstates of B, it is possible to construct eigenstates of B that are also (degenerate) eigenstates of A. In what follows, I carry out this construction.

We take advantage of the fact that any linear combination of the form

$$q = q_1 v(1) + q_2 v(2), \tag{42}$$

with arbitrary numbers q_1 and q_2, is an eigenstate of A corresponding to the degenerate eigenvalue 1. The question is: can we choose the numbers q_1 and q_2 so that the vector in Eq. 42 is also an eigenvector of B? In other words,

[1]This result might puzzle you if you expected $w(1)$ and $w(2)$ to be *the same* as $v(1)$ and $v(2)$. Because $v(1)$ and $v(2)$ are degenerate eigenvectors of A, *any* combination for the form $av(1) + bv(2)$ (where a and b are numbers) is also an eigenvector of A.

can we choose q_1 and q_2 so that

$$Bq = bq \tag{43}$$

where b is a number? Using Eq. 42 in Eq. 43 gives

$$Bq = q_1 Bv(1) + q_2 Bv(2) = b(q_1 v(1) + q_2 v(2)) \tag{44}$$

Now take the scalar product of this equation with $v(1)$. You get

$$q_1 v(1) \cdot Bv(1) + q_2 v(1) \cdot Bv(2) = b\, q_1 \tag{45}$$

I used the orthonormality of $\{v(1), v(2)\}$; that is, $v(1) \cdot v(2) = 0$ and $v(i) \cdot v(i) = 1$.

Take now the scalar product of Eq. 44 with $v(2)$ and perform the same simplifications. The result is

$$q_1 v(2) \cdot Bv(1) + q_2 v(2) \cdot Bv(2) = b\, q_2 \tag{46}$$

If we denote

$$Q_{ij} \equiv v(i) \cdot B \cdot v(j), \quad i, j = 1, 2 \tag{47}$$

then we can write Eqs. 45 and 46 as

$$\sum_{j=1}^{2} Q_{ij} q_j = b\, q_i \tag{48}$$

or, in matrix and vector notation,

$$Qq = b\, q \tag{49}$$

Here Q is the matrix with elements Q_{ij} and q is the vector with components q_i.

I can calculate Q_{ij} from Eq. 47. I did this in WorkBook 11 and found that

$$Q = \begin{pmatrix} 1 & -\mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix} \tag{50}$$

I calculated the eigenvalues of Q and found them to be 2 and 0, which are the same as the eigenvalues of B. I expected that to be the case, since the matrix Q is the matrix B in a representation using the basis set $\{v(1), v(2)\}$. The corresponding eigenvectors of Q are

$$q(1) = \{-\mathrm{i}, 1\} \text{ for eigenvalue } 0 \tag{51}$$

and

$$q(2) = \{\mathrm{i}, 1\} \text{ for eigenvalue } 2 \tag{52}$$

Because Q is Hermitian, these eigenvectors are orthogonal.

We started by forcing the vector

$$q = q_1 v(1) + q_2 v(2), \tag{53}$$

to satisfy the eigenvalue equation for B. This is how we derived equations for q_1 and q_2. We found two solutions: $\{-\mathrm{i}, 1\}$ and $\{\mathrm{i}, 1\}$. Inserting these into Eq. 53 gives

$$\begin{aligned} q(1) &\equiv -\mathrm{i}v(1) + v(2) \\ q(2) &\equiv \mathrm{i}v(1) + v(2) \end{aligned}$$

These were constructed to be eigenvectors of B. I verified in WorkBook 11 that they satisfy

$$Bq(1) = 0q(1)$$

and

$$Bq(2) = 2q(2)$$

Since $q(1)$ and $q(2)$ are linear combinations of $v(1)$ and $v(2)$ (which are the degenerate eigenvectors of A), they remain degenerate eigenvectors of A. It should be obvious why but if you do not see this, you can verify it.

The vectors $q(1)$ and $q(2)$ are not normalized but they can be normalized by the usual method. Then they will be orthonormal eigenvectors of both A and B, just as the theorem requires.

There is nothing in this method of construction that is specific to a two-dimensional vector space. We can do the same thing if A has a hundred degenerate eigenvalues and the dimension of the sapce is 300. The essence of this method is that we use the degenerate eigenvectors of A as a basis set to solve the eigenvalue problem for B.

§ 7 *Physical interpretation.* The theorem we just proved is central to the theory of measurement in quantum physics. Consider the case when $\hat{A}$ and $\hat{B}$ represent two observables A and B, and the operators commute. This means that we can find a complete orthonormal set of kets $|x_i\rangle$ such that

$$\hat{A}|x_i\rangle = \alpha_i|x_i\rangle \tag{54}$$

$$\hat{B}|x_i\rangle = \beta_i|x_i\rangle \tag{55}$$

This is not a bad notation but it is more informative to use the notation

$$|\alpha_i, \beta_j\rangle \tag{56}$$

for the common kets, with the consequence that Eqs. 54–55 are rewritten as

$$\hat{A}|\alpha_i, \beta_j\rangle = \alpha_i|\alpha_i, \beta_j\rangle \tag{57}$$

$$\hat{B}|\alpha_i, \beta_j\rangle = \beta_j|\alpha_i, \beta_j\rangle \tag{58}$$

The notation $|\alpha_i, \beta_j\rangle$ tells us that if the system is in this state and we measure A, we are sure to obtain α_i, and if we measure B, we are sure to obtain β_j. $|\alpha_i, \beta_j\rangle$ is *a pure state of both A and B.*

If a third operator $\hat{C}$ commutes with both $\hat{A}$ and $\hat{B}$, then we can find a complete orthonormal set of kets $|\alpha_i, \beta_j, \gamma_k\rangle$ for which

$$\hat{A}|\alpha_i, \beta_j, \gamma_k\rangle = \alpha_i|\alpha_i, \beta_j, \gamma_k\rangle \tag{59}$$

$$\hat{B}|\alpha_i, \beta_j, \gamma_k\rangle = \beta_j|\alpha_i, \beta_j, \gamma_k\rangle \tag{60}$$

and

$$\hat{C}|\alpha_i, \beta_j, \gamma_k\rangle = \gamma_k|\alpha_i, \beta_j, \gamma_k\rangle \tag{61}$$

This notation is of the same sort as $|n, \ell, m\rangle$ used for the hydrogen atom.

Suppose that for this particular physical system, there are only three observables A, B, and C, whose operators commute. In such a case, we say that A, B, and C form a *complete set of compatible observables.* For the hydrogen atom, this set is $\hat{H}$, $\hat{L}^2$, and $\hat{L}_z$ and the ket is $|n, \ell, m\rangle$. This set of observables is complete only because we ignore electron and nuclear spin.

Imagine that you use one of many numerical methods for solving the eigenvalue equation (the Schrödinger equation) for the hydrogen atom. *You do not take advantage of spherical symmetry* and program the computer to give the eigenstates $\psi_n(x, y, z)$ in Cartesian coordinates. The computer returns the eigenstates $\psi_1(x, y, z)$, $\psi_2(x, y, z)$, ... and the associated eigenvalues E_1, E_2, ... The numerical values of E_1, E_2, ... are the same as the ones given by Eq. 1. This must be so, because these energies are properties of the hydrogen atom and are independent of the method used to solve the equa-

tion. All correct numerical methods will give the same result for the energy eigenvalues.

Will the eigenstates $\psi_1(x,y,z)$, $\psi_2(x,y,z)$, ... be the same for all methods? The ground state is not degenerate. Therefore, if normalized, $\psi_1(x,y,z)$ must be the same as the analytical solution $\langle x,y,z\,|\,n=0,\ell=0,m=0\rangle$. It will not have the neat form $R_{0,0}(r)Y_0^0(\theta,\phi)$ provided by the analytical solution in polar coordinates, but for any point $\{x,y,z\}$, you can calculate r,θ,ϕ and then $R_{0,0}(r)Y_0^0(\theta,\phi)$ and obtain a number equal to $\psi_1(x,y,z)$.

Life gets more interesting when we examine $\psi_2(x,y,z)$, $\psi_3(x,y,z)$, $\psi_4(x,y,z)$, and $\psi_5(x,y,z)$. These energy eigenstates *are degenerate*: they all correspond to the energy given by Eq. 1 for $n=2$. However, if you compare $\psi_2(x,y,z)$, $\psi_3(x,y,z)$, $\psi_4(x,y,z)$, $\psi_5(x,y,z)$ to the analytical solutions $\langle x,y,z\,|\,n=1,\ell=0,m=0\rangle$, $\langle x,y,z\,|\,n=1,\ell=1,m=-1\rangle$, $\langle x,y,z\,|\,n=1,\ell=1,m=0\rangle$, $\langle x,y,z\,|\,n=1,\ell=1,m=1\rangle$, you *find that they are not the same.* Note also that the states $|n,\ell,m\rangle$ have more physical information than do $\psi_i(x,y,z)$, $i=2,\ldots,5$. $|n,\ell,m\rangle$ tells us that in this state the angular momentum squared has the value $\ell(\ell+1)\hbar^2$ and its projection on the z-axis is $m\hbar$. $\psi_i(x,y,z)$, $i=2,\ldots,5$, tells us nothing of the sort. The state $|n,\ell,m\rangle$ tells us what causes the degeneracy: in the state $|n,0,0\rangle$, the rotational energy is zero; in the states $|n,1,-1\rangle$, $|n,1,0\rangle$, and $|n,1,1\rangle$, the rotational energy is $\hbar^2 1(1+1)$ but the projections of the angular momentum on the OZ axis are different. When we get the results $\psi_2(x,y,z)$, $\psi_3(x,y,z)$, ... from the computer, we have no way of knowing the reason for the degeneracy. Nor are these states pure states of $\hat{L}^2$ and $\hat{L}_z$.

At this point we remember our theorem. Because $\hat{H}$ commutes with

$\hat{L}^2$ and $\hat{L}_z$, *we can construct* a set of eigenstates that are common to all three operators. We start with $\psi_i(x,y,z)$, $i = 2,\ldots,5$. Because these states are degenerate, they might not be orthonormal. We use the Gram-Schmidt procedure to produce orthonormal eigenstates $\phi_i(x,y,z)$, $i = 2,\ldots,5$. These are still degenerate eigenstates of $\hat{H}$. Then, we construct the four-dimensional matrix with elements $\langle \phi_i \,|\, \hat{L}^2 \,|\, \phi_j \rangle$ and find its four eigenvectors

$$x(i) = \{x_1(i), x_2(i), x_3(i), x_4(i)\},\ i = 1,2,3,4$$

The states

$$\chi_i(x,y,z) = \sum_{k=1}^{4} x_k(i)\phi_k(x,y,z)$$

are *eigenstates of energy and of* $\hat{L}^2$. One corresponds to eigenvalue 0 and the other three to eigenvalue $1(1+1)\hbar^2$. The latter three states are degenerate.

The fact that $\hat{L}^2$ is triply degenerate tells us that there must be another observable that commutes with $\hat{H}$ and $\hat{L}^2$. This is $\hat{L}_z$. We can now repeat the procedure: (1) Orthonormalize the three degenerate states, χ_i, to obtain a set $\lambda_i(x,y,z)$, $i = 2,3,4$; (2) calculate the matrix elements $\langle \lambda_i \,|\, \hat{L}_z \,|\, \lambda_j \rangle$ and find its three eigenvectors $y(i) = \{y_1(i), y_2(i), y_3(i)\}$, $i = 2,3,4$. Then

$$\eta_j(x,y,z) = \sum_j y_j(i)\lambda_i(x,y,z),\ i = 2,3,4$$

are eigenstates of $\hat{L}_z$, $\hat{L}^2$, and $\hat{H}$.

If you normalize χ_1, η_2, η_3, η_4 and compare them to $\langle x,y,z \,|\, n,\ell,m \rangle$ for $n = 2$, and $\ell = 0,1$, and $m = -\ell,\ldots,\ell$, you will find them to be numerically identical (except perhaps for a phase factor, which we can always take equal to 1). The theorem guided us to construct the states $\langle x,y,z \,|\, n,\ell,m \rangle$ from

the numerical results $\psi_1, \ldots, \psi_5$. The states $\langle x, y, z \,|\, n, \ell, m\rangle$ give us more physical insight into the properties of the atom than do the states $\psi_i(x, y, z)$.

Why was it that when you learned about the hydrogen atom you did not need to do so much work? Because your predecessors were clever. They took advantage of the spherical symmetry of the Hamiltonian to write $\hat{H}$ in a way that explicitly contained $\hat{L}^2$. Then they used the orthonormal eigenstates $Y_\ell^m(\theta, \phi)$ of $\hat{L}^2$ and looked for a solution of the Schrödinger equation in the form $R_{n,\ell}(r)Y_\ell^m(\theta, \phi)$. This resulted in an equation for the radial function $R_{n,\ell}(r)$ that they solved. The states $Y_\ell^m(\theta, \phi)$ were constructed to also be eigenstates of $\hat{L}_z$. Since you were just given the results, you were not privy to the method of construction.

This is all fine, but very confusing. So many states are possible; which one will appear in a given experiment? To answer this question, one must do the following. In most experiments, a system is exposed to some external agent (e.g. light, an electron beam, a collision with a molecule) that causes a change in its state. To find this new state, we must write down the Hamiltonian describing the experiment: the energy of the system plus the energy of the external agent plus the energy of the interaction between them. This Hamiltonian is used to write the time-dependent Schrödinger equation. Solving it gives the state $\psi(x, y, z)$ of the system created by its interaction with the external agent. For a hydrogen atom, we can analyze this state by using

$$\sum_n \sum_\ell \sum_m |n, \ell, m\rangle \langle n, \ell, m| = \hat{I}$$

and writing

$$\langle x, y, z \,|\, \psi\rangle = \sum_n \sum_\ell \sum_m \langle x, y, z \,|\, n, \ell, m\rangle \langle n, \ell, m \,|\, \psi\rangle$$

The probability that the system in state $|\psi\rangle$ has the energy E_n, the square of the angular momentum $\ell(\ell+1)\hbar^2$, and projection on the z-axis $m\hbar$ is given by

$$P_{n,\ell,m} = |\langle n, \ell, m \,|\, \psi\rangle|^2$$

Knowing these probabilities helps us understand how the system evolves when it is in the state $|\psi\rangle$, which is telling us how it is affected by the action of the external agent.

We could have expanded $|\psi\rangle$ by using the states ψ_2, ψ_3, ψ_4, ψ_5 mentioned earlier. However, these states tell us what the energy is but not the value of $\hat{L}^2$ or $\hat{L}_z$. Finding that the system is in the state $|\psi_2\rangle$, say, with a probability $|\langle\psi_2 \,|\, \psi\rangle|^2$ is not as informative as knowing $P_{n,\ell,m}$.

Section 11.2. Observables Whose Operators Do Not Commute: The Uncertainty Principle

§ 8 *Uncertainty in measurements.* We are interested in making measurements on two observables A and B. They are represented by the operators $\hat{A}$ and $\hat{B}$ whose spectra are $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$, respectively. The operators do not commute and the pure states of $\hat{A}$, denoted by $|a_n\rangle$, $n = 1, 2, \ldots$, are distinct from the pure states $|b_n\rangle$, $n = 1, 2, \ldots$ of $\hat{B}$. The system is in the state $|\psi\rangle$, which is not a pure state of either $\hat{A}$ or $\hat{B}$. Therefore we are uncertain what result we will get if we measure A or B. We ask here whether there is any relationship between the uncertainty we have when we measure A and the uncertainty we have when we measure B.

At first glance, this question makes no sense. If $\hat{A}$ and $\hat{B}$ are independent, why would the uncertainty in their measurements be related? On further

consideration, however, the question is a sensible one. If operators commute, then in principle there are states in which both observables that the operators represent have well defined values. What happens if the commutator is not zero but it is very small? In that case the two operators do not have common eigenstates but we expect that their eigenstates are close to each other. This means that if in some state we know precisely the value of one observable, we also know "almost precisely" the value of the other. This suggests that there might be a connection between the accuracy with which we can know A and B and the magnitude of the commutator $[\hat{A}, \hat{B}]$. In this section we set out to find that connection. The question is: if the state of the system is $|\psi\rangle$ and we contemplate measuring either A or B, is there any relationship between the uncertainty of our knowledge of the value of A and that of the value of B?

§ 9 *The magnitude of the uncertainty.* If we want to study uncertainty, we need to define a quantity whose magnitude tells us how large the uncertainty is. Let us assume that we study a quantity X and we know the probability $P(x)$ that X takes the value x. Probability theory tells us that a measure of the uncertainty in our knowledge of the outcome of the measurement of X is the standard deviation

$$\sigma = \sqrt{\langle (x - \langle x \rangle)^2 \rangle} \tag{62}$$

Here $\langle x \rangle$ means the average of x, defined as follows. If X takes discrete values x_1, x_2, ... then

$$\langle x \rangle \equiv \sum_{i=1}^{\infty} x_i P(x_i) \tag{63}$$

where $P(x_i)$ is the probability that X takes the value x_i. Also,

$$\langle x^2 \rangle \equiv \sum_{i=1}^{\infty} x_i^2 P(x_i) \tag{64}$$

The standard deviation σ is a measure of how different x could be from its average value $\langle x \rangle$. If it is probable that x takes values that are very different from the average value $\langle x \rangle$, then our ability to guess *a priori* the result of measuring X is diminished and σ is larger. The larger σ is, the less certain we are what the value of X would be when measured.

Exercise 5 Show that

$$\langle (x - \langle x \rangle)^2 \rangle = \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \tag{65}$$

As an example, consider experiments in which you throw a die. If the die is fair, the probability of getting a specific result $i \in \{1, 2, 3, 4, 5, 6\}$ in any one throw is

$$P(i) = \frac{1}{6} \tag{66}$$

The average value of the numbers produced by many throws is

$$\langle i \rangle = \sum_{i=1}^{6} iP(i) = \frac{1}{6} + \frac{2}{6} + \cdots + \frac{6}{6} = \frac{7}{2} \tag{67}$$

The standard deviation for fair-die throwing is

$$\begin{aligned} \sigma &= \sqrt{\sum_{i=1}^{6} (i - \langle i \rangle)^2 P(i)} \\ &= \sqrt{\left(1 - \frac{7}{2}\right)^2 \frac{1}{6} + \left(2 - \frac{7}{2}\right)^2 \frac{1}{6} + \cdots + \left(6 - \frac{7}{2}\right)^2 \frac{1}{6}} \\ &= \sqrt{\frac{35}{12}} = 1.71 \end{aligned} \tag{68}$$

But suppose that a crook has rigged the die so that it will never give 1 or 6. The probabilities are now

$$P(1) = P(6) = 0, \quad P(2) = P(3) = P(4) = P(5) = \frac{1}{4} \tag{69}$$

The average value is the same as for a fair die:

$$\langle i \rangle = \sum_{i=1}^{6} iP(i) = \frac{7}{2} \tag{70}$$

If we only measure the average value $\langle i \rangle$, we could not tell that the die is rigged. The standard deviation is

$$\sigma = \sqrt{\sum_{i=1}^{6} (i - \langle i \rangle)^2 P(i)} = \frac{\sqrt{5}}{2} = 1.12 \tag{71}$$

The standard deviation of a fair die is 1.71 and it is larger than that of a rigged one (which is 1.12); the crook has better information regarding the result of a throw.

Let us look at the case in which the die is rigged so that it is guaranteed to give 3. Then $P(1) = P(2) = P(4) = P(5) = P(6) = 0$ and $P(3) = 1$. It is easy to calculate that $\langle i \rangle = 3$ and $\sigma = 0$: we are *certain* of the result and the standard deviation is zero. Note that this rigged die is similar to a quantum system in the pure state $|3\rangle$; we know for certain that the outcome of a measurement (a throw) will be 3.

§ 10 *Standard deviation for pure and coherent states.* Let us use the concept of standard deviation to study the uncertainty with which we know the result of a measurement when we know the state of the system. We start with a system in a pure state $|a_n\rangle$ of the observable A. There is no uncertainty as to what the result of measuring A will be. It will be a_n, and if σ measures uncertainty then σ should be zero. Let us check that this is so.

The probability that a measurement of A gives a_m (given that the system is in the state $|a_n\rangle$) is

$$P_{a_n}(a_m) = |\langle a_m \,|\, a_n\rangle|^2 = \langle a_n \,|\, a_m\rangle \, \langle a_m \,|\, a_n\rangle = \delta_{nm} \tag{72}$$

So, $P_{a_n}(a_n) = 1$, and $P_{a_n}(a_m) = 0$ if $m \neq n$. The average value $\langle A\rangle_{a_n}$ of A, in this state, is

$$\langle A\rangle_{a_n} = \sum_{m=1}^{\infty} a_m P_{a_n}(a_m) = \sum_{m=1}^{\infty} a_m \delta_{nm} = a_n \tag{73}$$

It is very easy to calculate that

$$\sigma^2_{a_n}(A) \equiv \sum_{m=1}^{\infty} (a_m - \langle A\rangle_{a_n})^2 P_{a_n}(a_m) = \sum_{m=1}^{\infty} (a_m - a_n)^2 \delta_{nm} = 0 \tag{74}$$

The standard deviation is zero; this is just what we expected.

Now let us look at the case when the state $|\psi\rangle$ of the system is a coherent superposition of pure states $|a_m\rangle$. The probability of obtaining a_n in a measurement is

$$P_\psi(a_n) = |\langle a_n \,|\, \psi\rangle|^2 = \langle \psi \,|\, a_n\rangle \, \langle a_n \,|\, \psi\rangle \tag{75}$$

The average value of A is

$$\langle A\rangle_\psi = \sum_{n=1}^{\infty} a_n P_\psi(a_n) \tag{76}$$

The square of the standard deviation of A in state $|\psi\rangle$ is

$$\begin{aligned} \sigma^2_\psi(A) &= \sum_{n=1}^{\infty} (a_n - \langle A\rangle_\psi)^2 \, P_\psi(a_n) = \sum_{n=1}^{\infty} P_\psi(a_n) \left[a_n^2 - 2a_n \langle A\rangle_\psi + \langle A\rangle_\psi^2\right] \\ &= \langle A^2\rangle_\psi - \langle A\rangle_\psi^2 \end{aligned} \tag{77}$$

As long as $|\psi\rangle$ is a coherent superposition of pure states of A, $\sigma_\psi(A)$ cannot be zero.

Note that σ depends on both the state $|\psi\rangle$ and the observable A. Therefore, we cannot speak of the uncertainty of momentum or position without indicating *the state* of the system when the measurement is made. This is why I use the notation $\sigma_\psi(A)$ rather than $\sigma(A)$.

Exercise 6 The matrix

$$A = \begin{pmatrix} 1 & 0.7 & 7.3 \\ 0.7 & 2 & 2.1 \\ 7.3 & 2.1 & 3 \end{pmatrix}$$

represents an observable in the orthonormal basis set $|\alpha_1\rangle$, $|\alpha_2\rangle$, $|\alpha_3\rangle$ (that is, $A_{11} = \langle\alpha_1 \,|\, \hat{A} \,|\, \alpha_1\rangle$, $A_{12} = \langle\alpha_1 \,|\, \hat{A} \,|\, \alpha_2\rangle$, etc.) The matrix

$$B = \begin{pmatrix} 2.4 & 6.1 & 4.2 \\ 6.1 & 3.2 & 3.7 \\ 4.2 & 3.7 & 6 \end{pmatrix}$$

represents another observable in the same basis set.

1. Find the eigenvectors $v(1) = \{v_1(1), v_2(1), v_3(1)\}$, $v(2) = \{v_1(2), v_2(2), v_3(2)\}$, $v(3) = \{v_1(3), v_2(3), v_3(3)\}$ of A. The eigenket corresponding to the vector $v(1)$ is

   $$|v(1)\rangle = v_1(1)|\alpha_1\rangle + v_2(1)|\alpha_2\rangle + v_3(1)|\alpha_3\rangle$$

   and similarly for $|v(2)\rangle$ and $|v(3)\rangle$.

2. Calculate the standard deviation for A and B when the system is in the state $|v(1)\rangle$, in the state $|v(2)\rangle$, and in the state $|v(3)\rangle$.

3. Find the eigenvectors $w(1)$, $w(2)$, $w(3)$ of B and calculate the standard deviation for A when the system is in the state $|w(1)\rangle$, in the state $|w(2)\rangle$, and in the state $|w(3)\rangle$.

Section 11.3. The Uncertainty of Position and Momentum for a Particle Moving Freely in Space

The uncertainty principle was discovered by Heisenberg while analyzing the measurement of the position and momentum. We present here an analysis that is close to the one used by him. The general theory will be presented in Section 11.4.

§ 11 *The state in momentum representation.* We study a particle that moves through space with no force acting on it. Classically, the particle has a uniform velocity v and the momentum $p = mv$. I assume that the motion is one-dimensional; this simplifies the notation with no loss of insight.

The momentum in quantum mechanics is

$$p = \hbar k, \quad k \in [-\infty, +\infty] \tag{78}$$

where k is the *wave vector*. The ket $|k\rangle$ denotes a pure state in which a measurement of momentum is guaranteed to yield the value $p = \hbar k$. It is impossible to prepare in the laboratory a pure state $|k\rangle$; there will always be an uncertainty in the momentum. We can however prepare the coherent superposition

$$|\eta\rangle \equiv \int_{-\infty}^{+\infty} C \exp\left[-\frac{k^2}{a^2}\right] |k\rangle \, \mathrm{d}k \tag{79}$$

The expression $C\exp[-k^2/a^2]$ differs from zero only if k is not much larger than a. Therefore a is a measure of the uncertainty of momentum when the system is in state $|\eta\rangle$. The constant C will be determined from the requirement that to be physically meaningful, $|\eta\rangle$ must be normalized ($\langle\eta\,|\,\eta\rangle = 1$).

Because a function of the form $\exp[-x^2]$ is called a Gaussian, the ket $|\eta\rangle$ is called a *Gaussian wave packet in momentum representation.*

§ 12 *The uncertainty principle.* If the system is prepared in the state $|\eta\rangle$, the result of a measurement of the momentum is uncertain. We can calculate the standard deviation $\sigma_\eta(p)$ in such a measurement, to provide a quantitative description of this uncertainty. The result of a position measurement, when the system is in the state $|\eta\rangle$, is also uncertain. Here we ask the question: is there any relationship between the position uncertainty $\sigma_\eta(x)$ and the momentum uncertainty $\sigma_\eta(p)$? We will show that

$$\sigma_\eta(x)\sigma_\eta(p) = \hbar/2 \tag{80}$$

This equation tells us that if $\sigma_\eta(p)$ is large then $\sigma_\eta(x)$ must be small, and vice versa. This is Heisenberg's uncertainty principle.

To prove Eq. 80 we need to evaluate $\sigma_\eta(x)$ and $\sigma_\eta(p)$. This is easily done once we know $P_\eta(x)$ and $P_\eta(p)$.

§ 13 *Calculate $P_\eta(k)$ and $\sigma_\eta(p)$.* The probability amplitude that the wave vector of a particle in the state $|\eta\rangle$ has the value k' is

$$\begin{aligned}
\langle k'\,|\,\eta\rangle &= \int_{-\infty}^{+\infty} \mathrm{d}k\; C\exp\left[-\frac{k^2}{a^2}\right]\langle k'\,|\,k\rangle \\
&= \int_{-\infty}^{+\infty} \mathrm{d}k\; C\exp\left[-\frac{k^2}{a^2}\right]\delta(k'-k) \\
&= C\exp\left[-\frac{(k')^2}{a^2}\right]
\end{aligned} \tag{81}$$

In the first step I used $\langle k' \,|\, k\rangle = \delta(k'-k)$ and in the second, a known property of the δ-function.

The probability distribution that the particle has the wave vector k when the system is in state $|\eta\rangle$ is

$$P_\eta(k) = |\langle k \,|\, \eta\rangle|^2 = C^2 \exp\left[-\frac{2k^2}{a^2}\right] \tag{82}$$

The probability that the wave vector has a value between $-\infty$ and ∞ is equal to 1:

$$\int_{-\infty}^{+\infty} P_\eta(k)\,\mathrm{d}k = 1 \tag{83}$$

From this condition we obtain[2]

$$C = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \tag{84}$$

With this value for C, the Gaussian wave packet becomes

$$|\eta\rangle = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \exp\left[-\frac{k^2}{a^2}\right] |k\rangle \;\mathrm{d}k \tag{85}$$

and the probability distribution of k is

$$P_\eta(k) = \sqrt{\frac{2}{\pi}}\left(\frac{1}{a}\right) \exp\left[-\frac{2k^2}{a^2}\right] \tag{86}$$

A plot of this function versus k has a peak at $k = 0$ and tapers off rapidly when $|k| \gg a$. The value of a determines which values of the wave vector are very probable. Roughly, values of k between $-a$ and a are more probable than the others.

[2]All the integrals in this section were calculated by using `Mathematica`, in the file `WorkBook11.nb`.

The average value of the momentum is

$$\langle p \rangle_\eta = \int_{-\infty}^{+\infty} \hbar k P_\eta(k)\, \mathrm{d}k = 0 \tag{87}$$

This is not surprising: $P_\eta(k)$ is symmetric around $k = 0$ and therefore the probability that the particle has momentum $\hbar k$ is equal to the probability that it has momentum $-\hbar k$.

The average momentum squared is

$$\langle p^2 \rangle_\eta = \int_{-\infty}^{+\infty} (\hbar k)^2 P_\eta(k)\, \mathrm{d}k = \frac{a^2 \hbar^2}{4} \tag{88}$$

The standard deviation is

$$\sigma_\eta(p) = \sqrt{\langle p^2 \rangle_\eta - \langle p \rangle_\eta^2} = \frac{\hbar a}{2} \tag{89}$$

As expected, $\sigma_\eta(p)$ increases with a.

§ 14 *Calculate $P_\eta(x)$ and $\sigma_\eta(x)$.* The probability amplitude that a particle in the state $|\eta\rangle$ is located at position x is given by

$$\langle x \,|\, \eta \rangle = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \exp\left[-\frac{k^2}{a^2}\right] \langle x \,|\, k \rangle \, \mathrm{d}k \tag{90}$$

Here $|x\rangle$ is a pure state of the coordinate operator. The quantity $\langle x \,|\, k \rangle$ is given by

$$\langle x \,|\, k \rangle = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}kx}, \quad k \in [-\infty, +\infty] \tag{91}$$

Inserting Eq. 91 in Eq. 90 gives the Gaussian wave packet in coordinate representation,

$$\langle x \,|\, \eta \rangle = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \exp\left[-\frac{k^2}{a^2}\right] \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}kx} \, \mathrm{d}k \tag{92}$$

The integral was calculated in Cell 2, Section 3 of WorkBook 11.

The probability distribution is

$$P_\eta(x) = |\langle x \,|\, \eta\rangle|^2 = \frac{a}{\sqrt{2\pi}} \exp\left[-\frac{a^2x^2}{2}\right] \tag{93}$$

The likely values of the coordinate x for which $P_\eta(x)$ is not zero are those for which $a^2x^2/2 < 2$, or

$$-\frac{2}{a} < x < \frac{2}{a} \tag{94}$$

Note that the range of values of x that can be observed in a position measurement shrinks when a is large (see Eq. 93), and the range of the probable values of momentum grows when a increases (see Eq. 86). If the state $|\eta\rangle$ is such that we know momentum precisely, our knowledge of the position deteriorates! The analysis that follows expresses this statement in terms of the standard deviations $\sigma_\eta(x)$ and $\sigma_\eta(p)$.

We can now follow the recipe of quantum mechanics and calculate $\langle x\rangle_\eta$ and $\langle x^2\rangle_\eta$.

$$\langle x\rangle_\eta = \int_{-\infty}^{+\infty} P_\eta(x) x \,\mathrm{d}x = 0 \tag{95}$$

I performed the integral in Cell 3, Section 3 of WorkBook 11. Similarly,

$$\langle x^2\rangle_\eta = \int_{-\infty}^{+\infty} P_\eta(x) x^2 \,\mathrm{d}x = \frac{1}{a^2} \tag{96}$$

The standard deviation of the coordinate when the state is a Gaussian wave packet is (see WorkBook 11)

$$\sigma_\eta(x) = \sqrt{\langle x^2\rangle_\eta - \langle x\rangle_\eta^2} = \frac{1}{a} \tag{97}$$

§ 15 *Heisenberg's uncertainty principle.* The product of $\sigma_\eta(x)$ with $\sigma_\eta(p)$ is

$$\sigma_\eta(p)\sigma_\eta(x) = \frac{\hbar a}{2} \cdot \frac{1}{a} = \frac{\hbar}{2} \tag{98}$$

This is the famous Heisenberg uncertainty principle: in the state $|\eta\rangle$, the standard deviation of the position is inversely proportional to the standard deviation of the momentum. If the state is such that the momentum is known precisely, then the information we have about position is very poor. If we know the momentum exactly, then $\sigma_\eta(p) = 0$ and $\sigma_\eta(x) = \infty$; the particle can be located anywhere in space and we don't know where it is.

Heisenberg's paper played a very important role in the development of quantum mechanics. Prior to it, people hoped that one might be able to describe quantum dynamics through trajectories, which meant that one could know simultaneously, at any given time, the position and the momentum. Heisenberg put that notion to rest, forever.

§ 16 *A historical note.* In the paper in which the uncertainty principle was derived, Heisenberg gave a physical explanation for it. To observe the position of a particle in classical physics, we have to look at it. This means that we send photons towards the particle and observe the photons scattered by it. Optics has established that if we need a high spatial resolution in this observation, we must use light of small wavelength. But small wavelength means photons of high energy. When they bounce off a small particle, such as an electron or an atom, they give it a kick that changes its momentum. The more precise the position measurement, the harder the kick and the more the momentum of the particle changes. One can use the theory of the resolving power of a microscope to turn this argument into an equation similar to Eq. 98.

The theory of the resolving power of a microscope is not part of the obligatory knowledge of a theoretical physicist. At Heisenberg's Ph.D. ex-

amination, Professor Wien asked him to explain the resolving power of the microscope and the young genius answered with absolute silence. Heisenberg had been a poor student in Wien's laboratory course and Wien declared that Heisenberg could not receive a doctorate in physics. Sommerfeld, who was Heisenberg's supervisor, was aghast. After extensive negotiations with Wien, they agreed to give him the lowest passing grade. One of the greatest physicists of the 20th century barely got a Ph.D. in physics. After this unpleasant episode, Heisenberg read the theory of the microscope and used it a few years later to establish one of the most profound and unexpected rules of quantum mechanics.

The proof of Heisenberg's uncertainty relation given here made no reference to any specific method of measurement. The conclusion is restricted to measurements of momentum and position when the state is a Gaussian wave packet. For those of you who are interested in mathematics, it should be clear that Heisenberg's uncertainty principle follows from a property of Fourier transforms. If a function $f(x)$ has a shape similar to a bell curve, we can attribute to it a width a. We can then prove that the Fourier transform of f has width $1/a$. A pulse of an oscillating electric current is subject to an uncertainty principle. If we know the duration accurately, the frequency is uncertain.

Section 11.4. A General Theorem About Uncertainty in Measurements

§ 17 *Introduction.* Heisenberg's paper dealt with the connection between the uncertainty in our knowledge of position and of momentum, when the state of the system was a Gaussian wave packet. Our next question is whether

this relation is an exclusive property of measurements of position and momentum, or is more general. In previous sections I hinted that this property has to do with the fact that the operators of position and momentum do not commute. This suggests that perhaps a similar relation should exist between any pair of non-commuting operators.

In this section I derive such a relation, which was established by Robinson.[3] The plan is to:

1. State the theorem;
2. Examine its physical consequences;
3. Summarize a number of preliminary results (lemmas), which are proved in the Appendix;
4. Give the proof.

§ 18 *Robinson's uncertainty relation.* Before I state the theorem, we need to agree on what we mean by uncertainty. Suppose $\hat{A}$ and $\hat{B}$ are operators representing observables A and B, and the system is in a state $|\psi\rangle$ that is not an eigenstate of either $\hat{A}$ or $\hat{B}$. The average values of these operators, in this state, are

$$\langle A\rangle_\psi \equiv \langle\psi\,|\,\hat{A}\,|\,\psi\rangle \text{ and } \langle B\rangle_\psi \equiv \langle\psi\,|\,\hat{B}\,|\,\psi\rangle \tag{99}$$

We denote the deviation of $\hat{A}$ from its average value by $\delta\hat{A}^2$, and use a similar notation for $\hat{B}$; we have

$$\delta\hat{A}^2 \equiv (\hat{A} - \langle A\rangle_\psi)^2 \text{ and } \delta\hat{B}^2 \equiv (\hat{B} - \langle B\rangle_\psi)^2 \tag{100}$$

[3]H. P. Robinson, The uncertainty principle, *Physical Review 34*, 163-164 (1929)

When the system is in the state $|\psi\rangle$ these deviations have the mean values

$$\sigma_\psi(A)^2 \equiv \langle\psi\,|\,\delta\hat{A}^2\,|\,\psi\rangle = \langle\psi\,|\,\hat{A}^2\,|\,\psi\rangle - \langle A\rangle_\psi^2 \tag{101}$$

$$\sigma_\psi(B)^2 \equiv \langle\psi\,|\,\delta\hat{B}^2\,|\,\psi\rangle = \langle\psi\,|\,\hat{B}^2\,|\,\psi\rangle - \langle B\rangle_\psi^2 \tag{102}$$

These definitions generalize to operators the concept of mean-square deviation.

I can now state the theorem.

Let A and B be two observables, represented by the Hermitian operators $\hat{A}$ and $\hat{B}$, defined in the same space, *which do not commute.* Let $|\psi\rangle$ be a state that is not an eigenstate of either $\hat{A}$ or $\hat{B}$.

Under these conditions we have

$$\sigma_\psi(A)^2\sigma_\psi(B)^2 \geq \frac{1}{4}\left(|\langle\psi\,|\,[\hat{A},\hat{B}]\,|\,\psi\rangle|^2 + |\langle\psi\,|\,\{\delta\hat{A},\delta\hat{B}\}\,|\,\psi\rangle|^2\right) \tag{103}$$

where

$$[\hat{A},\hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \tag{104}$$

is the commutator of $\hat{A}$ and $\hat{B}$, and

$$\{\hat{A},\hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A} \tag{105}$$

is the *anticommutator* of $\hat{A}$ and $\hat{B}$.

Since the second term in the right-hand side of Eq. 103 is positive, the inequality

$$\sigma_\psi(A)^2\,\sigma_\psi(B)^2 \geq \frac{1}{4}|\langle\psi\,|\,[\hat{A},\hat{B}]\,|\,\psi\rangle|^2 \tag{106}$$

is also valid; this is a weaker condition than Eq. 103.

Before proving the theorem I will examine how it works for some examples to see what its physical consequences are.

§ 19 *Is the general formula consistent with Heisenberg's?* We can now check if the Heisenberg uncertainty formula, derived in Section 11.3, is consistent with (the more general statement) Eq. 103:

$$\sigma_\psi(A)^2\sigma_\psi(B)^2 \geq \frac{1}{4}|\langle\psi\,|\,[\hat{A},\hat{B}]\,|\,\psi\rangle|^2 + \frac{1}{4}|\langle\psi\,|\,\{\delta\hat{A},\delta\hat{B}\}\,|\,\psi\rangle|^2 \tag{107}$$

We want to see what this equation gives if A is momentum, B is the coordinate, and $|\psi\rangle$ is a Gaussian wave packet:

$$|\psi\rangle = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \exp\left[-\frac{k^2}{a^2}\right] |k\rangle \;\mathrm{d}k \tag{108}$$

Because $\langle\hat{x}\rangle_\eta$ and $\langle\hat{p}\rangle_\eta$ are zero, $\delta\hat{p} = \hat{p}$ and $\delta\hat{x} = \hat{x}$. I will use next the following known equations:

$$\langle x\,|\,k\rangle = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{ikx} \tag{109}$$

and

$$\langle x\,|\,\hat{p}\,|\,\psi\rangle = \frac{i}{\hbar}\frac{\partial}{\partial x}\langle x\,|\,k\rangle \tag{110}$$

Then I convert the commutator and the anticommutator to the coordinate representation and use Eqs. 109 and 110 in Eq. 107 to obtain

$$\langle\psi\,|\,[\hat{p},\hat{x}]\,|\,\psi\rangle = \int \mathrm{d}x\,\langle\psi\,|\,x\rangle\left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}x\,\langle x\,|\,\psi\rangle - x\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\langle x\,|\,\psi\rangle\right) = -\mathrm{i}\hbar \tag{111}$$

$$\langle\psi\,|\,\{\hat{p},\hat{x}\}\,|\,\rangle = \int \mathrm{d}x\,\langle\psi\,|\,x\rangle\left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}x\,\langle x\,|\,\psi\rangle + x\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\langle x\,|\,\psi\rangle\right) = 0 \tag{112}$$

All calculations needed for the evaluation of the integrands and of the integrals were performed in Cell 5, Section 3 of `Mathematica` file `WorkBook11.nb`. Using these two equations (Eq. 111 and Eq. 112) in the right-hand side of Eq. 103 gives (remember that $\delta\hat{p} = \hat{p}$ and $\delta\hat{x} = \hat{x}$)

$$\frac{1}{4}|\langle\eta\,|\,[\hat{p},\hat{x}]\,|\,\eta\rangle|^2 + \frac{1}{4}|\langle\eta\,|\,\{\hat{p},\hat{x}\}\,|\,\eta\rangle|^2 = \frac{\hbar^2}{4} \tag{113}$$

Using this in Eq. 107 gives

$$\sigma_\eta(p)^2\,\sigma_\eta(x)^2 = \frac{\hbar^2}{4} \tag{114}$$

We conclude that the general theorem, applied to $\hat{A} = \hat{p}$ and $\hat{B} = \hat{x}$ and a Gaussian wave-packet state $|\psi\rangle$, gives the same result as Eq. 98.

Note that a similar result holds *for any state* $|\psi\rangle$ *that is not an eigenstate of* $\hat{p}$ *or* $\hat{x}$, because we have $[\hat{x},\hat{p}] = \mathrm{i}\hbar$ and Eq. 106 becomes (for $\hat{A} = \hat{x}$ and $\hat{B} = \hat{p}$)

$$\sigma_\psi(x)^2\,\sigma_\psi(p)^2 \geq \frac{1}{4}|\langle\psi\,|\,\mathrm{i}\hbar\,|\,\psi\rangle|^2 = \frac{\hbar^2}{4} \tag{115}$$

This result is more general than Heisenberg's, which was limited to Gaussian wave packets.

§ 20 *Application to angular momentum.* I examine next the case when the state $|\psi\rangle = |l,m\rangle$ is an eigenstate of the square of the angular momentum and of the projection of the angular momentum $\hat{L}_z$. I want examine the accuracy I can obtain if I want to measure simultaneously the projections $\hat{L}_x$ and $\hat{L}_y$ of the angular momentum when the system is in a pure state of $\hat{L}_z$. I use Eq. 106 to calculate the mean square deviation in such a measurement. The right-hand side of this equation depends on the commutator of $\hat{L}_x$ with

$\hat{L}_y$, which has the value (see Metiu, *Quantum Mechanics*)

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z \tag{116}$$

Using all these in Eq. 106 gives

$$\sigma^2_{l,m}(\hat{L}_x)\,\sigma^2_{l,m}(\hat{L}_y) \geq \frac{1}{4}\langle l, m \,|\, \hat{L}_z \,|\, l, m\rangle = \frac{1}{4}\hbar m \tag{117}$$

The more accurately I measure $\hat{L}_x$, less accuracy I have in measuring $\hat{L}_y$ (if the system is in an eigenstate of $\hat{L}_z$). A similar relation holds regardless of what the state is, as long as it is not an eigenstate of $\hat{L}_x$ or $\hat{L}_y$.

§ 21 *Schwarz inequality.* Before I prove the theorem I need to collect or derive a number of results that will be used by the proof. One of them is the Schwarz inequality

$$\langle\psi \,|\, \psi\rangle\,\langle\phi \,|\, \phi\rangle \geq |\langle\psi \,|\, \phi\rangle|^2 \tag{118}$$

which holds for any pair of kets $|\psi\rangle$ and $|\phi\rangle$. Equality holds if one of the kets is zero or if one ket is a multiple (by a complex number) of the other. Since the theorem is an inequality, this relationship is a possible starting point for proving the theorem.

§ 22 *Some operator relationships.* I will use a few equations containing commutators, anticommutators and Hermitian conjugates. I give them without proof because you can derive them yourself. For operators $\hat{A}, \hat{B}, \hat{C}$ defined on the same space

$$\hat{A}\hat{B} = \frac{1}{2}[\hat{A}, \hat{B}] + \frac{1}{2}\{\hat{A}, \hat{B}\} \tag{119}$$

$$[\hat{A}, \alpha\hat{I}] = 0 \tag{120}$$

$$[\hat{A} - \alpha\hat{I}, \hat{B} - \beta\hat{I}] = [\hat{A}, \hat{B}] \tag{121}$$

where α and β are complex numbers.

§ 23 *Some properties related to Hermitian conjugation.* An *anti-Hermitian* operator is an operator that satisfies

$$\hat{O}^\dagger = -\hat{O} \tag{122}$$

If $\hat{A}$ and $\hat{B}$ are Hermitian operators defined on the same space then (1) the commutator is an anti-Hermitian operator:

$$[\hat{A}, \hat{B}]^\dagger = -[\hat{A}, \hat{B}]; \tag{123}$$

and (2) the anti-commutator is a Hermitian operator:

$$\{\hat{A}, \hat{B}\}^\dagger = \{\hat{A}, \hat{B}\} \tag{124}$$

While a Hermitian operator is equal to its Hermitian conjugate ($\hat{A}^\dagger = \hat{A}$), an anti-Hermitian operator is equal to the negative of its Hermitian conjugate. An analogy is helpful here. The Hermitian conjugation is a generalization of complex conjugation, a Hermitian operator is a generalization of a real number, and an anti-Hermitian operator is a generalization of an imaginary number.

§ 24 *A lemma concerning the matrix elements of Hermitian and anti-Hermitian operators.* In what follows we are going to use a Lemma concerning the matrix elements of hermitian and anti-Hermitian operators.
Lemma. Let $= |\psi\rangle$ *be an arbitrary ket,* $\hat{A}$ *a Hermitian operator, and* $\hat{B}$ *an anti-Hermitian operator. Then*

$$\langle\psi\,|\,\hat{A}\,|\,\psi\rangle \text{ is a real number} \tag{125}$$

and

$$\langle\psi\,|\,\hat{B}\,|\,\psi\rangle \text{ is an imaginary number} \tag{126}$$

It is not difficult to prove this Lemma. If an operator is defined by

$$\hat{O} \equiv \sum_{i=1}^{\infty} |\, o_i\rangle o_i \langle o_i \,| \tag{127}$$

then its adjoint operator is defined as (see Chapter 4)

$$\hat{O}^\dagger \equiv \sum_{i=1}^{\infty} |\, o_i\rangle o_i^* \langle o_i \,| \tag{128}$$

An anti-Hermitian operator satisfies, by definition, the relationship

$$\hat{O}^\dagger = -\hat{O} \tag{129}$$

It is possible to satisfy both Eqs. 129 and 128 only if the eigenvalues o_i satisfy $o_i^* = -o_i$. This means that the eigenvalues of an anti-Hermitian operator are imaginary numbers.

Now let us look at the matrix elements of one such operator:

$$\langle \psi \,|\, \hat{O} \,|\, \psi \rangle = \sum_{i=1}^{\infty} \langle \psi \,|\, o_i \rangle \, o_i \, \langle o_i \,|\, \psi \rangle \tag{130}$$

The expression $\langle \psi \,|\, o_i \rangle \, \langle o_i \,|\, \psi \rangle$ is a real number because $\langle o_i \,|\, \psi \rangle = \langle \psi \,|\, o_i \rangle^*$. Therefore, Eq. 130 tells us that if o_i is real, which is the case for Hermitian operators, then the matrix element $\langle \psi \,|\, \hat{O} \,|\, \psi \rangle$ is real. If o_i is imaginary, which is the case if $\hat{O}$ is anti-Hermitian, then the matrix element $\langle \psi \,|\, \hat{O} \,|\, \psi \rangle$ is imaginary. This completes the proof of the Lemma.

§ 25 *The proof of the theorem.* We have completed now our list of preliminary results and we can move on to prove the theorem. Define

$$\delta\hat{A} = \hat{A} - \langle A \rangle_\psi \hat{I} \tag{131}$$

$$\delta\hat{B} = \hat{B} - \langle B \rangle_\psi \hat{I} \tag{132}$$

Apply the Schwarz inequality to the kets $\delta\hat{A}|\psi\rangle$ and $\delta\hat{B}|\psi\rangle$. This gives

$$\langle\delta\hat{A}\psi\,|\,\delta\hat{A}\psi\rangle\langle\delta\hat{B}\psi\,|\,\delta\hat{B}\psi\rangle \geq |\langle\delta\hat{A}\psi\,|\,\delta\hat{B}\psi\rangle|^2 \tag{133}$$

Because $\delta\hat{A}$ and $\delta\hat{B}$ are Hermitian, they satisfy (see Chapter 4)

$$\langle\hat{O}\phi\,|\,\psi\rangle = \langle\phi\,|\,\hat{O}\,|\,\psi\rangle, \text{ where } \hat{O} = \delta\hat{A} \text{ or } \delta\hat{B} \tag{134}$$

Use this to write Eq. 133 as

$$\langle\psi\,|\,(\delta\hat{A})^2\,|\,\psi\rangle\langle\psi\,|\,(\delta\hat{B})^2\,|\,\psi\rangle \geq |\langle\psi\,|\,\delta\hat{A}\delta\hat{B}\,|\,\psi\rangle|^2 \tag{135}$$

By definition (see Eqs. 101–102 for the standard deviations)

$$\langle\psi\,|\,(\delta\hat{A})^2\,|\,\psi\rangle \equiv \sigma_\psi(A)^2 \tag{136}$$

$$\langle\psi\,|\,(\delta\hat{B})^2\,|\,\psi\rangle \equiv \sigma_\psi(B)^2 \tag{137}$$

Using these in Eq. 135 gives

$$\sigma_\psi(A)^2\sigma_\psi(B)^2 \geq |\langle\psi\,|\,\delta\hat{A}\delta\hat{B}\,|\,\psi\rangle|^2 \tag{138}$$

Use Eq. 119 to rewrite the right-hand side of Eq. 138:

$$\begin{aligned} |\langle\psi\,|\,\delta\hat{A}\delta\hat{B}\,|\,\psi\rangle|^2 &= |\langle\psi\,|\,\left(\frac{1}{2}[\delta\hat{A},\delta\hat{B}] + \frac{1}{2}\{\delta\hat{A},\delta\hat{B}\}\right)\,|\,\psi\rangle|^2 \\ &= \frac{1}{4}|\langle\psi\,|\,[\delta\hat{A},\delta\hat{B}]\,|\,\psi\rangle + \langle\psi\,|\,\{\delta\hat{A},\delta\hat{B}\}\,|\,\psi\rangle|^2 \end{aligned} \tag{139}$$

Now use Eq. 139 in Eq. 138):

$$\sigma_\psi(A)^2\sigma_\psi(B)^2 \geq \frac{1}{4}|\langle\psi\,|\,[\delta\hat{A},\delta\hat{B}]\,|\,\psi\rangle + \langle\psi\,|\,\{\delta\hat{A},\delta\hat{B}\}\,|\,\psi\rangle|^2 \tag{140}$$

But because $[\hat{A}-\alpha\hat{I},\hat{B}-\beta\hat{I}] = [\hat{A},\hat{B}]$ (see Eq. 121), we have

$$[\delta\hat{A},\delta\hat{B}] = [\hat{A},\hat{B}] \tag{141}$$

and Eq. 140 becomes

$$\sigma_\psi(A)^2\sigma_\psi(B)^2 \geq \frac{1}{4}|\langle\psi\,|\,[\hat{A},\hat{B}]\,|\,\psi\rangle + \langle\psi\,|\,\{\delta\hat{A},\delta\hat{B}\}\,|\,\psi\rangle|^2 \qquad (142)$$

The operator $[\hat{A},\hat{B}]$ is anti-Hermitian, and our Lemma tells us that $\langle\psi\,|\,[\hat{A},\hat{B}]\,|\,\psi\rangle$ is an *imaginary* number. The operator $\{\delta\hat{A},\delta\hat{B}\}$ is Hermitian, so $\langle\psi\,|\,\{\delta\hat{A},\delta\hat{B}\}\,|\,\psi\rangle$ is a *real* number.

Because of these properties, and the fact that $|a+ib|^2 = a^2+b^2$ when a and b are real numbers, the right-hand side of Eq. 142 simplifies to yield

$$\sigma_\psi(A)^2\,\sigma_\psi(B)^2 \geq \frac{1}{4}|\langle\psi\,|\,[\hat{A},\hat{B}]\,|\,\psi\rangle|^2 + \frac{1}{4}|\langle\psi\,|\,\{\delta\hat{A},\delta\hat{B}\}\,|\,\psi\rangle|^2 \qquad (143)$$

This is the inequality that we wanted to prove. Since the second term in the right-hand side is positive, the inequality

$$\sigma_\psi(A)^2\,\sigma_\psi(B)^2 \geq \frac{1}{4}|\langle\psi\,|\,[\hat{A},\hat{B}]\,|\,\psi\rangle|^2 \qquad (144)$$

is also true (but weaker than Eq. 143).

§ 26 *Summary.* In this chapter we have obtained several important results regarding our ability to know the result of measurements with certainty.

First, we have shown that if two operators $\hat{A}$ and $\hat{B}$ that represent observables A and B commute then

(a) all non-degenerate eigenstates of one operator are also eigenstates of the other operator;

(b) if one of the operators (say, $\hat{A}$) has n degenerate eigenstates corresponding to an eigenvalue α, these states might not be eigenstates of $\hat{B}$. However, they can be used to construct n degenerate eigenstates of $\hat{A}$ that are also eigenstates of $\hat{B}$.

This is important for several reasons. (1) If $\hat{A}$ and $\hat{B}$ commute, their nondegenerate states are pure states of both observables A and B. In such a state we know with certainty the magnitude of both A and B. (2) If the states are degenerate, we can construct states that are pure states of both operators. Thus, if $\hat{A}$ and $\hat{B}$ commute, it is possible to have states in which both observables have well defined values.

Since operators that commute have special properties with regard to measurements of the observables they represent, what can we say about measuring quantities whose operators do not commute? If the operators do not commute, they do not have joint eigenstates. There is no state in which we know with certainty the result of measuring *both* quantities.

To describe the uncertainty of a measurement of A, we use the standard deviation of A in state $|\psi\rangle$:

$$\sigma_\psi(A) = \sqrt{\langle A^2\rangle_\psi - \langle A\rangle_\psi^2}$$

We showed that for a particle whose free motion in space is described by a Gaussian wave packet $|\eta\rangle$, we have

$$\sigma_\eta(p)\,\sigma_\eta(x) = \frac{\hbar}{2}$$

This equation tells us that if the state $|\eta\rangle$ is such that we know the momentum with a high degree of certainty, then we have a great uncertainty in the value of position.

In a more general case, the uncertainties of measuring the values of two observables A and B, whose operators do not commute, are related through

$$\sigma_\psi(A)^2\,\sigma_\psi(B)^2 \geq \frac{1}{4}|\langle\psi\,|\,[\hat{A},\hat{B}]\,|\,\psi\rangle|^2 + \frac{1}{4}|\langle\psi\,|\,\{\delta\hat{A},\delta\hat{B}\}\,|\,\psi\rangle|^2$$

This seems to be a strange inequality, because $|\psi\rangle$, $\hat{A}$, and $\hat{B}$ have no physical relationship: A and B can be (almost) any observables and $|\psi\rangle$ can be (almost) any state. Nevertheless the inequality says that in any state $|\psi\rangle$ that is not an eigenstate of $\hat{A}$ or $\hat{B}$, knowing A with an accuracy $\sigma_\psi(A)$ puts limitations on the accuracy $\sigma_\psi(B)$ with which we can know B. Before you get mystical about this, remember that this statement is essentially the Schwarz inequality applied to states that are not pure states of either $\hat{A}$ or $\hat{B}$.

Note also that this statement about accuracy not only involves A and B but also depends importantly on $|\psi\rangle$. Statements like "if A is known accurately then B is not" are pointless unless we specify the state of the system.

Exercise 7 Calculate $\sigma_\psi(x)$, $\sigma_\psi(p)$, $\langle\psi\,|\,[\hat{p},\hat{x}]\,|\,\psi\rangle$, and $\langle\psi\,|\,\{\delta\hat{p},\delta\hat{x}\}\,|\,\psi\rangle$, and verify that

$$\sigma_\psi(x)^2\sigma_\psi(p)^2 \geq \frac{1}{4}|\langle\psi\,|\,[\hat{p},\hat{x}]\,|\,\psi\rangle|^2 + \frac{1}{4}|\langle\psi\,|\,\{\delta\hat{p},\delta\hat{x}\}\,|\,\psi\rangle|^2$$

for the following situations:

(a) $|\psi\rangle$ is the third excited state of a harmonic oscillator;

(b) $|\psi\rangle$ is the third excited state of a one-dimensional particle in a box.

Exercise 8 Suppose observables A and B correspond to operators $\hat{A}$ and $\hat{B}$ *that commute* and that $|\psi\rangle$ is not an eigenstate of either $\hat{A}$ or $\hat{B}$. Is there an

uncertainty in the values of A and B? Write a formula connecting $\sigma_\psi(A)$ and $\sigma_\psi(B)$. Calculate $\sigma_\psi(A)$ and $\sigma_\psi(B)$ for the case in which

$$|\psi\rangle = c_1|\alpha_1, \beta_1\rangle + c_2|\alpha_2, \beta_2\rangle$$

where

$$\hat{A}|\alpha_i, \beta_j\rangle = \alpha_i|\alpha_i, \beta_j\rangle$$

and

$$\hat{B}|\alpha_i, \beta_j\rangle = \beta_i|\alpha_i, \beta_j\rangle$$