

Chapter 9. Commutators and Measurements

Section 9.1. Observables Whose Operators Commute

§ 1 *Introduction.* A pure state of a hydrogen atom, in those cases in which we can ignore spin, is described 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*)

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

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

$$\hbar^2 \ell(\ell + 1) \quad (2)$$

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

$$\hbar m \quad (3)$$

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,

$$\hat{H}|n, \ell, m\rangle = E_n |n, \ell, m\rangle \quad (4)$$

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

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

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

$$P_{n,\ell,m}(x, y, z) dx dy dz = |\langle x, y, z | n, \ell, m \rangle|^2 dx dy dz \quad (7)$$

that the electron is inside the cube of volume $dx dy dz$ centered on the point $\{x, y, z\}$. The exact position is uncertain! One of the questions we answer in this section is why position is different from \hat{H} , \hat{L}^2 , or \hat{L}_z . Why can we have a situation in which we are certain of energy and the magnitudes of L^2 and L_z but uncertain of the position of the particle?

When we first determined the pure energy states of the hydrogen atom, we noted that \hat{H} , \hat{L}^2 , and \hat{L}_z commute with each other and suggested that this might be 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.

- (a) Consider two operators \hat{A} and \hat{B} , defined in the same space and representing observables. If all eigenstates of \hat{A} are eigenstates of \hat{B} and vice versa, then \hat{A} and \hat{B} commute.
- (b) If two operators, defined in the same space and representing observables, commute then:
 - (b1) All nondegenerate eigenstates of one operator are also eigenstates of the other operator.
 - (b2) If one operator, say \hat{A} , has n degenerate eigenstates, they are not necessarily eigenstates of \hat{B} . However, we can use the n degenerate eigenstates of \hat{A} to construct n orthonormal eigenstates of \hat{B} , which are also degenerate eigenstates of \hat{A} .

In other words, commuting operators have common eigenvectors, but we need to do some extra work to find them if degeneracy is present. We will spend quite a bit of time proving this theorem and understanding its meaning and consequences.

§ 3 *An example.* Let us examine a simple example to clarify why we insist that the operators must belong to the same space. It will also show why we say that if \hat{A} and \hat{B} commute, it is possible to find a common set of eigenstates and we do not say that all eigenstates of \hat{A} must be eigenstates of \hat{B} (or vice versa).

Consider a harmonic oscillator with the Hamiltonian

$$\hat{H} = a \frac{d^2}{dx^2} + bx^2 \quad (8)$$

where a and b are two constants whose values are of no interest here. Notice that this operator does not change if we make the transformation $x \rightarrow -x$.

We introduce next the operator $\hat{\sigma}$ defined by

$$\hat{\sigma}\psi(x) \equiv \psi(-x) \quad (9)$$

This operator is typical of the symmetry operators studied in group theory. We can see by direct calculation that $\hat{\sigma}$ commutes with \hat{H} :

$$\hat{H}\hat{\sigma}\psi(x) = a\frac{d^2}{dx^2}\psi(-x) + bx^2\psi(-x) \quad (10)$$

$$\begin{aligned} \hat{\sigma}\hat{H}\psi(x) &= \hat{\sigma}\left(a\frac{d^2}{dx^2}\psi(x) + bx^2\psi(x)\right) \\ &= a\frac{d^2}{dx^2}\psi(-x) + bx^2\psi(-x) \end{aligned} \quad (11)$$

We used the facts that $\hat{\sigma} d^2/dx^2 = d^2/dx^2$ and $\hat{\sigma}x^2 = x^2$. Since $\psi(x)$ is an arbitrary function, we conclude that

$$\hat{H}\hat{\sigma} = \hat{\sigma}\hat{H} \quad (12)$$

According to the theorem in §2, since the two operators commute, they can have common eigenvectors.

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

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

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.

Now for the eigenfunctions of $\hat{\sigma}$. *Any* function f having the property

$$f(-x) = -f(x) \quad (13)$$

(an anti-symmetric function) satisfies the eigenvalue equation

$$\hat{\sigma}f(x) = -f(x) \quad (14)$$

so f is an eigenvector of $\hat{\sigma}$ corresponding to the eigenvalue -1.

Exercise 1 Which of xe^{-x^2} , $x^3e^{-x^2}$, ..., $x^{2n+1}e^{-x^2}$, ... are eigenstates of $\hat{\sigma}$ with the eigenvalue -1?

Any function g with the property

$$g(-x) = g(x) \quad (15)$$

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

$$\hat{\sigma}g(x) = g(x) \quad (16)$$

Exercise 2 Show that $x^{2n}e^{-x^2}$ is an eigenstate of $\hat{\sigma}$ with the eigenvalue $+1$, for any integer value of n .

Obviously the eigenvalues of $\hat{\sigma}$ have extremely high degeneracy! All possible 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:

$$\begin{aligned} \hat{\sigma}\phi_n(x) &\equiv \phi_n(-x) = -\phi_n(x) && \text{if } n \text{ is odd} \\ \hat{\sigma}\phi_n(x) &\equiv \phi_n(-x) = \phi_n(x) && \text{if } n \text{ is even} \end{aligned} \quad (17)$$

Every eigenstate of \hat{H} is an eigenstate of $\hat{\sigma}$, as promised. This is in agreement with the theorem, which says that because the eigenstates of \hat{H} are not degenerate and $[\hat{H}, \hat{\sigma}] = 0$, they 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, \dots are eigenfunctions of $\hat{\sigma}$ with the eigenvalue $+1$. They are definitely *not* eigenfunctions of \hat{H} . What went wrong? It is here that the phrase “act on the same linear space” comes into play. The wave functions $\phi_n(x)$ of the harmonic oscillator must satisfy

$$\int_{-\infty}^{+\infty} \phi_n(x)^* \phi_n(x) dx = 1; \quad (18)$$

they belong to the space L^2 . The functions x^2, x^4, x^6, \dots do not satisfy Eq. 18; 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

$$\psi(x) = \sum_{n=0}^{\infty} c_n \phi_n(x) \quad (19)$$

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

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

$$s(x) \equiv \sum_{k=0}^{\infty} c_k \phi_{2k}(x) \quad (20)$$

The anti-symmetric functions are given by

$$a(x) \equiv \sum_{k=0}^{\infty} b_k \phi_{2k+1}(x) \quad (21)$$

Here c_k and b_k are 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 3 I claim that we must have $\sum_{k=0}^{\infty} b_k^* b_k < \infty$. Explain why I say that.

Obviously

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

and

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

so they 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 $\phi_n(x)$, $n = 0, 1, 2, \dots$. They are eigenstates of both operators and they form a complete orthonormal basis set for the space on which both operators act.

This example shows that, if not properly understood, the theorem can be grievously misused. We had trouble with $\hat{\sigma}$ but not with \hat{H} . Why is that? The eigenstates of \hat{H} are *not degenerate* while those of $\hat{\sigma}$ have infinite degeneracy. When we apply the theorem, degenerate states require additional care. You'll see soon how we handle them.

§ 4 *The states of the hydrogen atom.* For the hydrogen atom, the theorem does not seem to cause any trouble. The eigenstates $|n, \ell, m\rangle$ satisfy the theorem for the commuting operators \hat{H} , \hat{L}^2 , and \hat{L}_z , and the hydrogen atom seems to give us no chance of being wrong. But it does! It is true that 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 them to have these properties. Consider the state

$$|\psi_{n,\ell}\rangle = \sum_{m=-\ell}^{+\ell} c_m |n, \ell, m\rangle \quad (22)$$

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 but it is *not* an eigenstate of \hat{L}_z :

$$\hat{L}_z |\psi_{n,\ell}\rangle = \sum_{m=-\ell}^{+\ell} c_m \hbar m |n, \ell, m\rangle \neq \hbar m \sum_{m=-\ell}^{+\ell} c_m |n, \ell, m\rangle \quad (23)$$

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 4 Show that $|\psi\rangle$ defined by Eq. 22 is an eigenstate of \hat{H} and of \hat{L}^2 .

Exercise 5 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 ?

§ 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 if operators \hat{A} and \hat{B} have the same eigenstates, then they commute. Let $|u_i\rangle$, $i = 1, 2, \dots$, be the common eigenstates, with corresponding eigenvalues α_i for \hat{A} and β_i for \hat{B} . Then we can write

$$\hat{A} = \sum_{i=1}^{\infty} |u_i\rangle \alpha_i \langle u_i| \quad (24)$$

and

$$\hat{B} = \sum_{i=1}^{\infty} |u_i\rangle \beta_i \langle u_i| \quad (25)$$

We can then calculate:

$$\begin{aligned} \hat{A}\hat{B} &= \left(\sum_{i=1}^{\infty} |u_i\rangle \alpha_i \langle u_i| \right) \left(\sum_{j=1}^{\infty} |u_j\rangle \beta_j \langle u_j| \right) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |u_i\rangle \alpha_i \langle u_i | u_j \rangle \beta_j \langle u_j| \end{aligned} \quad (26)$$

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

$$\hat{A}\hat{B} = \sum_{i=1}^{\infty} |u_i\rangle \alpha_i \beta_i \langle u_i| \quad (27)$$

A similar calculation shows that $\hat{B}\hat{A} = \sum_{i=1}^{\infty} |u_i\rangle \beta_i \alpha_i \langle u_i| = \sum_{i=1}^{\infty} |u_i\rangle \alpha_i \beta_i \langle u_i|$. We have found that $\hat{A}\hat{B} = \hat{B}\hat{A}$ and therefore $[\hat{A}, \hat{B}] = 0$, which is what we wanted to prove.

Now let us assume that $[\hat{A}, \hat{B}] = 0$ and show that \hat{A} and \hat{B} have common eigenstates. Denote the eigenvalues of \hat{A} by α_i and the corresponding eigenvectors by $|u_i\rangle$; we have

$$\hat{A}|u_i\rangle = \alpha_i |u_i\rangle \quad (28)$$

and

$$\hat{A} = \sum_{i=1}^{\infty} |u_i\rangle \alpha_i \langle u_i| \quad (29)$$

I consider first the case when all the eigenvalues of \hat{A} are *nondegenerate*. Because $\hat{A}\hat{B} = \hat{B}\hat{A}$, we have, for each $|u_i\rangle$,

$$\hat{A}\hat{B}|u_i\rangle = \hat{B}\hat{A}|u_i\rangle = \alpha_i \hat{B}|u_i\rangle \quad (30)$$

This means that $\hat{B}|u_i\rangle$ is an eigenvector of \hat{A} corresponding to the eigenvalue α_i . Since this eigenvalue *is not degenerate*, $\hat{B}|u_i\rangle$ must be proportional to $|u_i\rangle$. If the proportionality constant is denoted by β_i , then we can write

$$\hat{B}|u_i\rangle = \beta_i|u_i\rangle \quad (31)$$

This means that the eigenstate $|u_i\rangle$ of \hat{A} is also an eigenstate of \hat{B} , which is what we intended to prove.

We started our analysis by looking at the eigenstates of \hat{A} . 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 an eigenstate $|u_i\rangle$ of \hat{A} *corresponds to a degenerate eigenvalue* α_i , this argument falls apart and we have to dig deeper. Let us denote by $|u_{i,1}\rangle, |u_{i,2}\rangle, \dots, |u_{i,n_i}\rangle$ all linearly independent eigenvectors of \hat{A} corresponding to the eigenvalue α_i . Since $\hat{A}\hat{B} = \hat{B}\hat{A}$,

$$\hat{A}\hat{B}|u_{i,j}\rangle = \hat{B}\hat{A}|u_{i,j}\rangle = \alpha_i\hat{B}|u_{i,j}\rangle, \quad j = 1, 2, \dots, n_i \quad (32)$$

so all n_i states $\hat{B}|u_{i,j}\rangle, j = 1, 2, \dots, n_i$, are eigenvectors of \hat{A} corresponding to the eigenvalue α_i . Because of this, each $\hat{B}|u_{i,j}\rangle$ is a linear combination of the states $|u_{i,m}\rangle$:

$$\hat{B}|u_{i,j}\rangle = \sum_{m=1}^{n_i} c_{j,m} |u_{i,m}\rangle, \quad j = 1, 2, \dots, n_i \quad (33)$$

We can readily verify that $\hat{B}|u_{i,j}\rangle$ given by Eq. 33 satisfies the eigenvalue equation for \hat{A} :

$$\hat{A}\hat{B}|u_{i,j}\rangle = \sum_{m=1}^{n_i} c_{j,m} \hat{A}|u_{i,m}\rangle = \alpha_i \sum_{m=1}^{n_i} c_{j,m} |u_{i,m}\rangle = \alpha_i \hat{B}|u_{i,j}\rangle \quad (34)$$

Note that the sum in Eq. 33 is *confined to the eigenvectors of \hat{A} corresponding to α_i* . Eq. 33 tells us that $\{|u_{i,j}\rangle\}_{j=1}^{n_i}$ forms a complete basis set for representing all $\hat{B}|u_{i,j}\rangle, j = 1, \dots, n_i$. By using the Gram-Schmidt procedure, we can convert this basis set to an orthonormal one. The resulting collection of orthonormal kets are linear combinations of $\{\hat{B}|u_{i,j}\rangle\}_{j=1}^{n_i}$ and therefore are eigenstates of \hat{A} corresponding to the eigenvalue α_i . Next, we use the theorem that assures us that there is a unitary operator \hat{U} that

converts this orthonormal basis set into a set of orthonormal eigenvectors $\{|x_{i,j}\rangle\}_{j=1}^{n_i}$ of \hat{B} . Since these orthonormal eigenvectors of \hat{B} are linear combinations of the set $\{|u_{i,j}\rangle\}_{j=1}^{n_i}$, they remain eigenvectors of \hat{A} corresponding to α_i . This proves our theorem: if \hat{A} and \hat{B} commute, we can construct a complete set of orthonormal vectors that are eigenvectors of both \hat{A} and \hat{B} , even when some of the eigenvectors are degenerate.

As the examples discussed earlier in this chapter show, the theorem does not mean that all degenerate eigenvectors of \hat{A} must be eigenvectors of \hat{B} , or vice versa.

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

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (35)$$

and

$$B = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \quad (36)$$

Obviously they are Hermitian. To save mental effort, I have done all the required calculations in the `Mathematica` file `WorkBook9.nb`.

To see if the matrices commute, I calculate

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \quad (37)$$

and

$$BA = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \quad (38)$$

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 eigenstates of A are (see `WorkBook9`)

$$\left. \begin{array}{l} av(1) = \{0, 1\} \quad \text{with eigenvalue } a_1 = 1 \\ av(2) = \{1, 0\} \quad \text{with eigenvalue } a_2 = 1 \end{array} \right\} \quad (39)$$

They are degenerate. The components of $av(j)$ are denoted by $av_i(j)$, $i = 1, 2$, so that, for example, $av_1(1) = 0$ and $av_2(1) = 1$.

The eigenstates and eigenvalues of B are

$$\left. \begin{aligned} bv(1) &= \{i, 1\} && \text{with eigenvalue } b_1 = 2 \\ bv(2) &= \{-i, 1\} && \text{with eigenvalue } b_2 = 0 \end{aligned} \right\} \quad (40)$$

The eigenstates $bv(1)$ and $bv(2)$ are not degenerate.

Part **b1** of the theorem says that, because the eigenstates of B are not degenerate, they must be eigenstates of A . This means that $A \cdot bv(i)$ is proportional to $bv(i)$, $i = 1, 2$, and the proportionality constants are the eigenvalues a_1 and a_2 . I tested this conclusion in the *Mathematica* file *WorkBook9* and found it to be correct.

Next I tested whether the degenerate eigenvectors of A are eigenvectors of B . They are not (see *WorkBook9*). 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, we will carry out this construction.

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

$$q = q_1 av(1) + q_2 av(2), \quad (41)$$

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

$$Bq = bq? \quad (42)$$

Using Eq. 41 in Eq. 42 gives

$$B \cdot q = q_1 B \cdot av(1) + q_2 B \cdot av(2) = b(q_1 av(1) + q_2 av(2)) \quad (43)$$

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

$$q_1 av(1) \cdot B \cdot av(1) + q_2 av(1) \cdot B \cdot av(2) = bq_1 \quad (44)$$

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

Take now the scalar product of Eq. 43 with $av(2)$ and perform the same simplifications. The result is

$$q_1 av(2) \cdot B \cdot av(1) + q_2 av(2) \cdot B \cdot av(2) = bq_2 \quad (45)$$

If we denote

$$Q_{ij} \equiv av(i) \cdot B \cdot av(j), \quad i, j = 1, 2 \quad (46)$$

then we can write Eqs. 44 and 45 as

$$\sum_{j=1}^2 Q_{ij} q_j = b q_i \quad (47)$$

or, in matrix and vector notation,

$$Qq = bq \quad (48)$$

Here Q is the matrix with elements Q_{ij} and q is the vector with components q_i . Note that in Eq. 41 we represent the vector q in terms of the basis set $\{av(1), av(2)\}$, and Q_{ij} given by Eq. 46 are the matrix elements of B in the same basis.

I can calculate Q_{ij} from Eq. 46. I did this in *WorkBook9* and found that

$$Q = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \quad (49)$$

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

$$q(1) = \{-i, 1\} \text{ for eigenvalue } 0 \quad (50)$$

and

$$q(2) = \{i, 1\} \text{ for eigenvalue } 2 \quad (51)$$

Because Q is Hermitian, these eigenvectors are orthogonal.

We can now go back to the original representation. The vectors

$$qo(i) = q_1(i) av(1) + q_2(i) av(2), \quad i = 1, 2 \quad (52)$$

must be eigenvectors of the matrix B . In *WorkBook9*, I evaluated Eq. 52 and obtained

$$\begin{aligned} qo(1) &= \{1, -i\} \\ qo(2) &= \{1, i\} \end{aligned}$$

I verified that they are eigenvectors of B . Therefore, $qo(1)$ and $qo(2)$ are eigenvectors of B that are also eigenvectors of A , corresponding to the degenerate eigenvalue. They are not normalized; normalization gives the eigenvectors (see Workbook9)

$$\left\{ \frac{1}{\sqrt{2}}, -\frac{i}{\sqrt{2}} \right\}$$

and

$$\left\{ \frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}} \right\}$$

§ 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

$$\hat{A}|x_i\rangle = \alpha_i|x_i\rangle \quad (53)$$

$$\hat{B}|x_i\rangle = \beta_i|x_i\rangle \quad (54)$$

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

$$|\alpha_i, \beta_j\rangle \quad (55)$$

for the common basis of kets, with the consequence that Eqs. 53–54 are rewritten as

$$\hat{A}|\alpha_i, \beta_j\rangle = \alpha_i|\alpha_i, \beta_j\rangle \quad (56)$$

$$\hat{B}|\alpha_i, \beta_j\rangle = \beta_j|\alpha_i, \beta_j\rangle \quad (57)$$

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

$$\hat{A}|\alpha_i, \beta_j, \gamma_k\rangle = \alpha_i|\alpha_i, \beta_j, \gamma_k\rangle \quad (58)$$

$$\hat{B}|\alpha_i, \beta_j, \gamma_k\rangle = \beta_j|\alpha_i, \beta_j, \gamma_k\rangle \quad (59)$$

and

$$\hat{C}|\alpha_i, \beta_j, \gamma_k\rangle = \gamma_k|\alpha_i, \beta_j, \gamma_k\rangle \quad (60)$$

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 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, \dots . The numerical values of E_1, E_2, \dots 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 equation. All correct numerical methods will give the same values for the 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 $\psi_1(x, y, z)$ must be the same as $\langle x, y, z | n = 0, \ell = 0, m = 0 \rangle$. 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 $\psi_1(x, y, z)$ at that point.

Life gets more interesting when we examine $\psi_2(x, y, z)$, $\psi_3(x, y, z)$, $\psi_4(x, y, z)$, $\psi_5(x, y, z)$. 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 $\langle x, y, z | n = 1, \ell = 0, m = 0 \rangle$, $\langle x, y, z | n = 1, \ell = 1, m = -1 \rangle$, $\langle x, y, z | n = 1, \ell = 1, m = 0 \rangle$, $\langle x, y, z | n = 1, \ell = 1, m = 1 \rangle$, you *find that they are not the same*. Note also that $|n, \ell, m\rangle$ has more physical information than do $\psi_i(x, y, z)$, $i = 2, \dots, 5$. $|n, \ell, m\rangle$ 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 $m\hbar$. $\psi_i(x, y, z)$, $i = 2, \dots, 5$, 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 $\psi_i(x, y, z)$, $i = 2, \dots, 5$, and use the Gram-Schmidt procedure to produce orthonormal eigenstates $\phi_i(x, y, z)$, $i = 2, \dots, 5$. Then, we construct the four-dimensional matrix with elements $\langle \phi_i | \hat{L}^2 | \phi_j \rangle$ and find its four eigenvectors

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

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 $1(1+1)\hbar^2$.

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, 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 $\mathbf{y}(i) = \{y_1(i), y_2(i), y_3(i)\}$, $i = 2, 3, 4$. Then

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

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

If you orthogonalize $\chi_1, \eta_2, \eta_3, \eta_4$ and compare them to $\langle x, y, z | n, \ell, m \rangle$ for $n = 2$, and $\ell = 0, 1$, and $m = -\ell, \dots, \ell$, you will find them to be identical (except perhaps for a phase factor, which we can always take equal to 1). The theorem guided us to construct the states $\langle x, y, z | n, \ell, m \rangle$, which give us more physical insight into the state of the atom than do the $\psi_i(x, y, z)$ -s.

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 use the orthonormal eigenstates $Y_\ell^m(\theta, \phi)$ of \hat{L}^2 to construct eigenstates of \hat{H} . 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 $\psi_2, \psi_3, \psi_4, \psi_5$ 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 9.2. Observables Whose Operators Do Not Commute: The Uncertainty Principle

§ 8 *Introduction.*

§ 9 *Uncertainty in measurements.* You have heard of the uncertainty principle, which says that certain observables are connected in a strange way: the more accurately you know the value of one of them, the less accurately you know the value of the other. The position of a particle and its momentum are one such pair: if the system is in a state for which you know exactly

where a particle is, then you have no information about how fast it is moving (all velocities are equally probable). Here I want to make these statements more precise, mathematically and physically.

When is there such uncertainty in quantum mechanics? As you have already learned, quantum mechanics deals with two kinds of states: pure states of an observable and coherent states.¹ Let us review these concepts. An observable A has a spectrum a_1, a_2, \dots , and $\alpha \in D$ consisting of all values that A can possibly take in an experiment. The spectrum of A is the same as the set of all eigenvalues of \hat{A} . The system can be prepared in one of the pure states $|a_1\rangle, |a_2\rangle, \dots, |\alpha\rangle$, $\alpha \in D$, of A ; these pure states are eigenstates of the operator \hat{A} representing the observable A . The pure states have a remarkable property: if a clever experimentalist manages to prepare the system in the pure state $|a_n\rangle$ and then measures the value of A , she is guaranteed to obtain a_n . There is no uncertainty regarding the result of the measurement.

What happens if she prepares the system in state $|a_n\rangle$ and then measures a property B ? B is an observable, which has a spectrum b_1, b_2, \dots , and $\beta \in D'$, and its pure states are $|b_1\rangle, |b_2\rangle, \dots, |\beta\rangle$, $\beta \in D'$. These are the eigenvalues and eigenvectors of the operator \hat{B} representing B . Quantum mechanics tells us that if the system is in the state $|a_n\rangle$ and B is measured, then the probability that the measurement of B gives the result b_m is

$$P_{a_n}(b_m) = |\langle b_m | a_n \rangle|^2 \quad (61)$$

$P_{a_n}(b_m)$ does not have to be 0 or 1. This means that, in general, when we are certain that the system is in state $|a_n\rangle$, we don't know the result we get when we measure B . We only know *the probability* that the value b_m is obtained. This is why we say that there is an uncertainty regarding the value of B .

This uncertainty can disappear in one and only one instance: when \hat{A} and \hat{B} commute. Then the two operators have joint eigenvectors:

$$\hat{A}|a_n, b_m\rangle = a_n|a_n, b_m\rangle \quad (62)$$

and

$$\hat{B}|a_n, b_m\rangle = b_m|a_n, b_m\rangle; \quad (63)$$

a pure state of A is also a pure state of B and vice versa. If an experimentalist prepares the system in the state $|a_n, b_m\rangle$ then he is certain that if he measures

¹Quantum statistical mechanics deals with a third kind of states, described by density matrices. These states are ignored in these lectures.

A the result will be a_n and that if he measures B the result will be b_m . No doubts here: in the state $|a_n, b_m\rangle$ the values of *both* A and B are known with certainty.

If the system is in a state $|\psi\rangle$ that is a coherent superposition of the pure states $|a_n\rangle$ of A,

$$|\psi\rangle = \sum_{i=1}^{\infty} |a_n\rangle c_n, \quad (64)$$

then the probability that a measurement of A yields the value a_m is

$$P_{\psi}(a_m) = |\langle a_m | \psi \rangle|^2 = |c_m|^2 \quad (65)$$

(I used $\langle a_m | a_n \rangle = \delta_{mn}$.) The outcome of the measurement is uncertain; its probability is, however, known.

The point of this analysis is that whether the result of a measurement is certain or not depends on the state of the system. If the state is an eigenstate of \hat{A} , the result of a measurement of A is certain. If the system is in a joint eigenstate of \hat{A} and \hat{B} , we are certain of the result of measuring A or B. If the state is not an eigenstate of \hat{A} , the result of measuring A is uncertain.

The uncertainty principle deals with a different situation. It asks the following question: if the state is $|\psi\rangle$ and I contemplate measuring either A or B, is there any relationship between the uncertainty of measuring A and that of measuring B?

§ 10 *The magnitude of the uncertainty.* We talk about uncertainty. How do we measure it? 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 is the standard deviation

$$\sigma = \sqrt{\langle (x - \langle x \rangle)^2 \rangle} \quad (66)$$

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

$$\langle x \rangle = \sum_{i=1}^{\infty} x_i P(x_i) \quad (67)$$

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

$$\langle x^2 \rangle = \sum_{i=1}^{\infty} x_i^2 P(x_i) \quad (68)$$

The standard deviation σ is the average “distance” between the magnitude of x and the average value $\langle x \rangle$. If x can take values that are very different from the average value $\langle x \rangle$, then our ability to guess the result 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 6 Show that

$$\langle (x - \langle x \rangle)^2 \rangle = \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \quad (69)$$

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

$$P(i) = \frac{1}{6} \quad (70)$$

The average value of the numbers produced by many throws is

$$\langle i \rangle = \sum_{i=1}^6 iP(i) = \frac{1}{6} + \frac{2}{6} + \cdots + \frac{6}{6} = \frac{7}{2} \quad (71)$$

The standard deviation for fair-die throwing is

$$\begin{aligned} \sigma &= \sqrt{\sum_{i=1}^6 (i - \langle i \rangle)^2 P(i)} \\ &= \sqrt{\left(1 - \frac{7}{2}\right)^2 \frac{1}{6} + \left(2 - \frac{7}{2}\right)^2 \frac{1}{6} + \cdots + \left(6 - \frac{7}{2}\right)^2 \frac{1}{6}} \\ &= \sqrt{\frac{35}{12}} = 1.71 \end{aligned} \quad (72)$$

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

$$P(1) = P(6) = 0, \quad P(2) = P(3) = P(4) = P(5) = \frac{1}{4} \quad (73)$$

The average value is

$$\langle i \rangle = \sum_{i=1}^6 iP(i) = \frac{7}{2} \quad (74)$$

The standard deviation is

$$\sigma = \sqrt{\sum_{i=1}^6 (i - \langle i \rangle)^2 P(i)} = \frac{\sqrt{5}}{2} = 1.12 \quad (75)$$

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 the die in this last example is in a “pure state” $|3\rangle$; we know for sure that the outcome of a measurement (a throw) will be 3.

§ 11 *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

$$P_{a_n}(a_m) = |\langle a_m | a_n \rangle|^2 = \langle a_n | a_m \rangle \langle a_m | a_n \rangle = \delta_{nm} \quad (76)$$

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

$$\langle A \rangle_{a_n} = \sum_{m=1}^{\infty} a_m P_{a_n}(a_m) = \sum_{m=1}^{\infty} a_m \delta_{nm} = a_n \quad (77)$$

It is very easy to calculate that

$$\sigma_{a_n}^2(A) \equiv \sum_{m=1}^{\infty} (a_m - \langle A \rangle_{a_n})^2 P_{a_n}(a_m) = \sum_{m=1}^{\infty} (a_m - a_n)^2 \delta_{nm} = 0 \quad (78)$$

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

$$P_\psi(a_n) = |\langle a_n | \psi \rangle|^2 = \langle \psi | a_n \rangle \langle a_n | \psi \rangle \quad (79)$$

The average value of A is

$$\langle A \rangle_\psi = \sum_{n=1}^{\infty} a_n P_\psi(a_n) \quad (80)$$

We have seen this formula before. The square of the standard deviation of A in state $|\psi\rangle$ is

$$\begin{aligned} \sigma_\psi^2(A) &= \sum_{n=1}^{\infty} (a_n - \langle A \rangle_\psi)^2 P_\psi(a_n) = \sum_{n=1}^{\infty} a_n^2 P_\psi(a_n) - \left(\sum_{n=1}^{\infty} a_n P_\psi(a_n) \right)^2 \\ &= \langle A^2 \rangle_\psi - \langle A \rangle_\psi^2 \end{aligned} \quad (81)$$

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 7 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 $\mathbf{v}(1) = \{v_1(1), v_2(1), v_3(1)\}$, $\mathbf{v}(2) = \{v_1(2), v_2(2), v_3(2)\}$, $\mathbf{v}(3) = \{v_1(3), v_2(3), v_3(3)\}$ of A . The eigenket corresponding to the vector $\vec{v}(1)$ 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 $\mathbf{w}(1)$, $\mathbf{w}(2)$, $\mathbf{w}(3)$ 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 9.3. 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 9.4.

§ 12 *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.

According to de Broglie, the momentum in quantum mechanics is

$$p = \hbar k, \quad k \in [-\infty, +\infty] \quad (82)$$

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 state

$$|\eta\rangle \equiv \int_{-\infty}^{+\infty} c_k |k\rangle dk \equiv \int_{-\infty}^{+\infty} C \exp\left[-\frac{k^2}{a^2}\right] |k\rangle dk \quad (83)$$

The coefficient $c_k \equiv C \exp[-k^2/a^2]$ differs from zero only when 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*.

§ 13 *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

$$\sigma_\eta(x)\sigma_\eta(p) = \hbar/2 \quad (84)$$

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

To prove Eq. 84 we need to evaluate $\sigma_\eta(x)$ and $\sigma_\eta(p)$. This is easily done once we know $P_\eta(x)$ and $P_\eta(p)$.

§ 14 *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

$$\begin{aligned} \langle k' | \eta \rangle &= \int_{-\infty}^{+\infty} dk C \exp\left[-\frac{k^2}{a^2}\right] \langle k' | k \rangle \\ &= \int_{-\infty}^{+\infty} dk C \exp\left[-\frac{k^2}{a^2}\right] \delta(k' - k) \\ &= C \exp\left[-\frac{(k')^2}{a^2}\right] \end{aligned} \quad (85)$$

In the first step I used $\langle k' | k \rangle = \delta(k' - k)$ and in the second, Eq. 2.30, which tells us how to deal with integrals containing δ -functions.

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

$$P_\eta(k) = |\langle k | \eta \rangle|^2 = C^2 \exp\left[-\frac{2k^2}{a^2}\right] \quad (86)$$

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

$$\int_{-\infty}^{+\infty} P_{\eta}(k) dk = 1 \quad (87)$$

From this condition we obtain²

$$C = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \quad (88)$$

With this value for C , the Gaussian wave packet becomes

$$|\eta\rangle = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \exp\left[-\frac{k^2}{a^2}\right] |k\rangle dk \quad (89)$$

and the probability distribution of k is

$$P_{\eta}(k) = \sqrt{\frac{2}{\pi}} \left(\frac{1}{a}\right) \exp\left[-\frac{2k^2}{a^2}\right] \quad (90)$$

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 what values the wave vector is likely to have (roughly between $-a$ and a).

The average value of the momentum is (see [WorkBook9](#))

$$\langle p \rangle_{\eta} = \int_{-\infty}^{+\infty} \hbar k P_{\eta}(k) dk = 0 \quad (91)$$

The average momentum squared is (see [WorkBook9](#))

$$\langle p^2 \rangle_{\eta} = \int_{-\infty}^{+\infty} (\hbar k)^2 P_{\eta}(k) dk = \frac{a^2 \hbar^2}{4} \quad (92)$$

The standard deviation is

$$\sigma_{\eta}(p) = \sqrt{\langle p^2 \rangle_{\eta} - \langle p \rangle_{\eta}^2} = \frac{\hbar a}{2} \quad (93)$$

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

²All the integrals in this section were calculated in the file [WorkBook9.nb](#).

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

$$\langle x | \eta \rangle = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \exp\left[-\frac{k^2}{a^2}\right] \langle x | k \rangle dk \quad (94)$$

Here $|x\rangle$ is a pure state of the coordinate operator. The quantity $\langle x | k \rangle$ is the Schrödinger eigenstate of the momentum operator (in general, $\langle x | \psi \rangle$ is the wave function in Schrödinger theory). You know that the momentum operator is $\hat{p} = (\hbar/i)(\partial/\partial x)$ (Metiu, *Quantum Mechanics*, p. 14) and therefore the momentum eigenvalue equation is

$$\frac{\hbar}{i} \frac{\partial}{\partial x} \langle x | k \rangle = p \langle x | k \rangle \quad (95)$$

The solution is

$$\langle x | k \rangle = A e^{-ikx}, \quad k \in [-\infty, +\infty] \quad (96)$$

where A is an unknown constant. You can check easily that Eq. 96 is a solution of Eq. 95 when

$$p = \hbar k \quad (97)$$

Inserting Eq. 96 in Eq. 94 gives for the Gaussian wave packet in coordinate representation

$$\begin{aligned} \langle x | \eta \rangle &= \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \exp\left[-\frac{k^2}{a^2}\right] A e^{-ikx} dk \\ &= A \sqrt{a} (2\pi)^{1/4} \exp\left[-\frac{a^2 x^2}{4}\right] \end{aligned} \quad (98)$$

The integral was calculated in Cell 2, Section 3 of **WorkBook9**.

The probability that the particle is between x and $x + dx$ is given by

$$P_\eta(x) dx = A^2 a (2\pi)^{1/2} \exp\left[-\frac{a^2 x^2}{2}\right] dx \quad (99)$$

The probability distribution must satisfy

$$\int_{-\infty}^{+\infty} P_\eta(x) dx = 1, \quad (100)$$

from which we can determine A (see [WorkBook9](#)):

$$A = \frac{1}{\sqrt{2\pi}} \quad (101)$$

With this value of A , the probability distribution is

$$P_\eta(x) = \frac{a}{\sqrt{2\pi}} \exp\left[-\frac{a^2 x^2}{2}\right] \quad (102)$$

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

$$-\frac{2}{a} < x < \frac{2}{a} \quad (103)$$

Note that the range of values of x that can be observed in a position measurement shrinks when a is large (see Eq. 102), and the range of the probable values of momentum grows when a increases (see Eq. 90). 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 $\langle x^2 \rangle_\eta$.

$$\langle x \rangle_\eta = \int_{-\infty}^{+\infty} P_\eta(x) x \, dx = 0 \quad (104)$$

I performed the integral in Cell 3, Section 3 of [WorkBook9](#). Similarly,

$$\langle x^2 \rangle_\eta = \int_{-\infty}^{+\infty} P_\eta(x) x^2 \, dx = \frac{1}{a^2} \quad (105)$$

The standard deviation of the coordinate when the state is a Gaussian wave packet is (see [WorkBook9](#))

$$\sigma_\eta(x) = \sqrt{\langle x^2 \rangle_\eta - \langle x \rangle_\eta^2} = \frac{1}{a} \quad (106)$$

§ 16 *The uncertainty principle.* The product of $\sigma_\eta(x)$ with $\sigma_\eta(p)$ is

$$\sigma_\eta(p)\sigma_\eta(x) = \frac{\hbar a}{2} \cdot \frac{1}{a} = \frac{\hbar}{2} \quad (107)$$

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.

In the paper in which the uncertainty principle was presented, 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. 107.

The proof given here is more general than this physical argument since it relies on the laws of quantum mechanics.

The theory of the resolving power of a microscope is not part of the obligatory knowledge of a theoretical physicist. At his Ph.D. examination, Professor Wien asked Heisenberg to explain the resolving power 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 degree 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.

Section 9.4. A General Theorem about Uncertainty

§ 17 *The uncertainty theorem.* We have followed Heisenberg and derived, in the previous section, the uncertainty “principle” for the coordinate and momentum. Our next question is whether there is a more general statement concerning the uncertainty of the measurements of two arbitrary observables. The answer is yes and it is given by the following 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} . The standard deviation $\sigma_\psi(A)$ in a measurement of A , when the system is in the state $|\psi\rangle$, is

$$\sigma_\psi(A)^2 = \langle \psi | (\hat{A} - \langle \hat{A} \rangle_\psi)^2 | \psi \rangle \equiv \langle \psi | \delta \hat{A} | \psi \rangle \quad (108)$$

Similarly

$$\sigma_\psi(B)^2 = \langle \psi | (\hat{B} - \langle \hat{B} \rangle_\psi)^2 | \psi \rangle \equiv \langle \psi | \delta \hat{B} | \psi \rangle \quad (109)$$

Here

$$\langle \hat{A} \rangle_\psi \equiv \langle \psi | \hat{A} | \psi \rangle \quad \text{and} \quad \langle \hat{B} \rangle_\psi \equiv \langle \psi | \hat{B} | \psi \rangle \quad (110)$$

are the average values of A and B when the system is in the state $|\psi\rangle$.

Under these conditions we have

$$\sigma_\psi(A)^2 \sigma_\psi(B)^2 \geq \frac{1}{4} \left(|\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle|^2 + |\langle \psi | \{\delta \hat{A}, \delta \hat{B}\} | \psi \rangle|^2 \right) \quad (111)$$

where

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \quad (112)$$

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

$$\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A} \quad (113)$$

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

Before I prove this theorem, I need to establish a few things. Recall that the Schwarz inequality is

$$\langle \psi | \psi \rangle \langle \phi | \phi \rangle \geq |\langle \psi | \phi \rangle|^2 \quad (114)$$

which holds for any kets $|\psi\rangle$ and $|\phi\rangle$. Equality holds if either one of the kets is zero or if one is a multiple (by a complex number) of the other.

It is easy to establish a number of relationships needed for the proof of the theorem, which I give as exercises.

Exercise 8 For operators \hat{A} , \hat{B} , \hat{C} defined on the same space, prove the following.

$$[\hat{A} + \hat{B}, \hat{C}] = [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}] \quad (115)$$

$$[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] \quad (116)$$

$$\hat{A}\hat{B} = \frac{1}{2}[\hat{A}, \hat{B}] + \frac{1}{2}\{\hat{A}, \hat{B}\} \quad (117)$$

$$[\hat{A}, \alpha\hat{I}] = 0 \quad (118)$$

$$[\hat{A}, \alpha\hat{B}] = \alpha[\hat{A}, \hat{B}] \quad (119)$$

$$[\hat{A} - \alpha\hat{I}, \hat{B} - \beta\hat{I}] = [\hat{A}, \hat{B}] \quad (120)$$

where α and β are complex numbers.

Exercise 9 Prove that if \hat{A} and \hat{B} are Hermitian operators defined on the same space then

1. $[\hat{A}, \hat{B}]^\dagger = -[\hat{A}, \hat{B}]$ (the commutator is an *anti-Hermitian* operator)
2. $\{\hat{A}, \hat{B}\}^\dagger = \{\hat{A}, \hat{B}\}$ (the anti-commutator is a Hermitian operator)

Now we prove the theorem. Define

$$\delta\hat{A} = \hat{A} - \langle A \rangle_\psi \hat{I} \quad (121)$$

$$\delta\hat{B} = \hat{B} - \langle B \rangle_\psi \hat{I} \quad (122)$$

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

$$\langle \delta\hat{A}\psi | \delta\hat{A}\psi \rangle \langle \delta\hat{B}\psi | \delta\hat{B}\psi \rangle \geq |\langle \delta\hat{A}\psi | \delta\hat{B}\psi \rangle|^2 \quad (123)$$

Because $\delta\hat{A}$ and $\delta\hat{B}$ are Hermitian, they satisfy

$$\langle \hat{O}\phi | \psi \rangle = \langle \phi | \hat{O}\psi \rangle, \text{ where } \hat{O} = \delta\hat{A} \text{ or } \delta\hat{B} \quad (124)$$

Use this to write Eq. 123 as

$$\langle \psi | (\delta \hat{A})^2 | \psi \rangle \langle \psi | (\delta \hat{B})^2 | \psi \rangle \geq |\langle \psi | \delta \hat{A} \delta \hat{B} | \psi \rangle|^2 \quad (125)$$

It is easy to see that

$$\langle \psi | (\delta \hat{A})^2 | \psi \rangle = \sigma_\psi(A)^2 \quad (126)$$

$$\langle \psi | (\delta \hat{B})^2 | \psi \rangle = \sigma_\psi(B)^2 \quad (127)$$

Use these in Eq. 125 to obtain

$$\sigma_\psi(A)^2 \sigma_\psi(B)^2 \geq |\langle \psi | \delta \hat{A} \delta \hat{B} | \psi \rangle|^2 \quad (128)$$

Now we use $\hat{O}\hat{P} = \frac{1}{2}[\hat{O}, \hat{P}] + \frac{1}{2}\{\hat{O}, \hat{P}\}$ to rewrite the right-hand side of Eq. 128:

$$\begin{aligned} |\langle \psi | \delta \hat{A} \delta \hat{B} | \psi \rangle|^2 &= |\langle \psi | \left(\frac{1}{2}[\delta \hat{A}, \delta \hat{B}] + \frac{1}{2}\{\delta \hat{A}, \delta \hat{B}\} \right) | \psi \rangle|^2 \\ &= \frac{1}{4} |\langle \psi | [\delta \hat{A}, \delta \hat{B}] | \psi \rangle + \langle \psi | \{\delta \hat{A}, \delta \hat{B}\} | \psi \rangle|^2 \end{aligned} \quad (129)$$

Therefore (use Eq. 129 in Eq. 128)

$$\sigma_\psi(A)^2 \sigma_\psi(B)^2 \geq \frac{1}{4} |\langle \psi | [\delta \hat{A}, \delta \hat{B}] | \psi \rangle + \langle \psi | \{\delta \hat{A}, \delta \hat{B}\} | \psi \rangle|^2 \quad (130)$$

But because $[\hat{A} - \alpha \hat{I}, \hat{B} - \beta \hat{I}] = [\hat{A}, \hat{B}]$, we have

$$[\delta \hat{A}, \delta \hat{B}] = [\hat{A}, \hat{B}] \quad (131)$$

and Eq. 130 becomes

$$\sigma_\psi(A)^2 \sigma_\psi(B)^2 \geq \frac{1}{4} |\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle + \langle \psi | \{\delta \hat{A}, \delta \hat{B}\} | \psi \rangle|^2 \quad (132)$$

The operator $[\delta \hat{A}, \delta \hat{B}]$ is anti-Hermitian, so $\langle \psi | [\delta \hat{A}, \delta \hat{B}] | \psi \rangle$ 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. 132 can be simplified:

$$\sigma_\psi(A)^2 \sigma_\psi(B)^2 \geq \frac{1}{4} |\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle|^2 + \frac{1}{4} |\langle \psi | \{\delta \hat{A}, \delta \hat{B}\} | \psi \rangle|^2 \quad (133)$$

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

$$\sigma_\psi(A)^2 \sigma_\psi(B)^2 \geq \frac{1}{4} |\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle|^2 \quad (134)$$

is also true (but weaker than Eq. 133).

§ 18 *Is the general formula consistent with Heisenberg's?* We can now check if the formula derived in Section 9.3 is consistent with the general uncertainty principle, Eq. 111:

$$\sigma_\psi(A)^2 \sigma_\psi(B)^2 \geq \frac{1}{4} |\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle|^2 + \frac{1}{4} |\langle \psi | \{\delta\hat{A}, \delta\hat{B}\} | \psi \rangle|^2 \quad (135)$$

We want to see what this equation gives if A is momentum, B is the coordinate, and $|\psi\rangle = |\eta\rangle$ where $|\eta\rangle$ was defined by Eq. 89.

Because $\langle \hat{x} \rangle_\eta$ and $\langle \hat{p} \rangle_\eta$ are zero, $\delta\hat{p} = \hat{p}$ and $\delta\hat{x} = \hat{x}$. In Cell 5, Section 3 of *WorkBook9*, I calculated that

$$\langle \eta | [\hat{p}, \hat{x}] | \eta \rangle = \int dx \langle \eta | x \rangle \left(\frac{\hbar}{i} \frac{\partial}{\partial x} x \langle x | \eta \rangle - x \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x | \eta \rangle \right) = -i\hbar \quad (136)$$

$$\langle \eta | \{\hat{p}, \hat{x}\} | \eta \rangle = \int dx \langle \eta | x \rangle \left(\frac{\hbar}{i} \frac{\partial}{\partial x} x \langle x | \eta \rangle + x \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x | \eta \rangle \right) = 0 \quad (137)$$

Using these two equations in the right-hand side of Eq. 111 gives (remember that $\delta\hat{p} = \hat{p}$ and $\delta\hat{x} = \hat{x}$)

$$\frac{1}{4} |\langle \eta | [\hat{p}, \hat{x}] | \eta \rangle|^2 + \frac{1}{4} |\langle \eta | \{\hat{p}, \hat{x}\} | \eta \rangle|^2 = \frac{\hbar^2}{4} \quad (138)$$

On the other hand, Eq. 107 gives

$$\sigma_\eta(p)^2 \sigma_\eta(x)^2 = \frac{\hbar^2}{4} \quad (139)$$

The general theorem, applied to $\hat{A} = \hat{p}$ and $\hat{B} = \hat{x}$ and a Gaussian wavepacket state $|\eta\rangle$ gives the same result as Eq. 107.

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 $[\hat{x}, \hat{p}] = i\hbar$ and Eq. 134 becomes (for $\hat{A} = \hat{x}$ and $\hat{B} = \hat{p}$)

$$\sigma_\psi(x)^2 \sigma_\psi(p)^2 \geq \frac{1}{4} |\langle \psi | i\hbar | \psi \rangle|^2 = \frac{\hbar^2}{4} \quad (140)$$

§ 19 *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, there are special states that are pure states of both operators. By using examples we have shown that these states have a more transparent physical meaning. Knowing them makes it easier to interpret the physical meaning of other states of the system.

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, therefore, 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} . The uncertainty is a result of the geometry of the space in which quantum mechanics operates.

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 10 For a Gaussian wave packet $|\eta\rangle$, show that the domain of p for which $P_\eta(p)$ is not zero increases as $\sigma_\eta(x)$ decreases and becomes narrower as $\sigma_\eta(x)$ increases. What is the analogous statement for $P_\eta(x)$?

Exercise 11 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 12 Suppose observables A and B correspond to operators \hat{A} and \hat{B} that commute and that $|\psi\rangle$ is not an eigenvalue 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

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