

UNCERTAINTY PRINCIPLE AND COMPATIBLE OBSERVABLES

B. Zwiebach
October 21, 2013

Contents

1	Uncertainty defined	1
2	The Uncertainty Principle	3
3	The Energy-Time uncertainty	6
4	Lower bounds for ground state energies	9
5	Diagonalization of Operators	11
6	The Spectral Theorem	12
7	Simultaneous Diagonalization of Hermitian Operators	16
8	Complete Set of Commuting Observables	18

1 Uncertainty defined

As we know, observables are associated to Hermitian operators. Given one such operator A we can use it to measure some property of the physical system, as represented by a state Ψ . If the state is in an eigenstate of the operator A , we have no uncertainty in the value of the observable, which coincides with the eigenvalue corresponding to the eigenstate. We only have uncertainty in the value of the observable if the physical state is not an eigenstate of A , but rather a superposition of various eigenstates with different eigenvalues.

We want to define the **uncertainty** $\Delta A(\Psi)$ of the Hermitian operator A on the state Ψ . This uncertainty should vanish if and only if the state is an eigenstate of A . The uncertainty, moreover, should be a real number. In order to define such uncertainty we first recall that the expectation value of A on the state Ψ , assumed to be normalized, is given by

$$\langle A \rangle = \langle \Psi | A | \Psi \rangle = \langle \Psi, A \Psi \rangle. \quad (1.1)$$

The expectation $\langle A \rangle$ is guaranteed to be real since A is Hermitian. We then define the uncertainty as the norm of the vector obtained by acting with $(A - \langle A \rangle I)$ on the physical state (I is the identity operator):

$$\Delta A(\Psi) \equiv \left| (A - \langle A \rangle I) \Psi \right|. \quad (1.2)$$

The uncertainty, so defined is manifestly non-negative. If the uncertainty is zero, the vector inside the norm is zero and therefore:

$$\Delta A(\Psi) = 0 \quad \rightarrow \quad (A - \langle A \rangle I)\Psi = 0 \quad \rightarrow \quad A\Psi = \langle A \rangle\Psi, \quad (1.3)$$

and the last equation confirms that the state is indeed an eigenstate of A (note that $\langle A \rangle$ is a number). You should also note that $\langle A \rangle$ is indeed the eigenvalue, since taking the eigenvalue equation $A\Psi = \lambda\Psi$ and forming the inner product with another Ψ we get

$$\langle \Psi, A\Psi \rangle = \lambda \langle \Psi, \Psi \rangle = \lambda \quad \rightarrow \quad \lambda = \langle A \rangle. \quad (1.4)$$

Alternatively, if the state Ψ is an eigenstate, we now know that the eigenvalue is $\langle A \rangle$ and therefore the state $(A - \langle A \rangle I)\Psi$ vanishes and its norm is zero. We have therefore shown that

The uncertainty $\Delta A(\Psi)$ vanishes if and only if Ψ is an eigenstate of A .

(1.5)

To compute the uncertainty one usually squares the expression in (1.2) so that

$$(\Delta A(\Psi))^2 = \langle (A - \langle A \rangle I)\Psi, (A - \langle A \rangle I)\Psi \rangle \quad (1.6)$$

Since the operator A is assumed to be Hermitian and consequently $\langle A \rangle$ is real, we have $(A - \langle A \rangle I)^\dagger = A - \langle A \rangle I$, and therefore we can move the operator on the first entry onto the second one to find

$$(\Delta A(\Psi))^2 = \langle \Psi, (A - \langle A \rangle I)^2 \Psi \rangle. \quad (1.7)$$

While this is a reasonable form, we can simplify it further by expansion

$$(\Delta A(\Psi))^2 = \langle \Psi, (A^2 - 2\langle A \rangle A + \langle A \rangle^2 I)\Psi \rangle. \quad (1.8)$$

The last two terms combine and we find

$(\Delta A(\Psi))^2 = \langle A^2 \rangle - \langle A \rangle^2.$

(1.9)

Since the left-hand side is greater than or equal to zero, this incidentally shows that the expectation value of A^2 is larger than the expectation value of A , squared:

$\langle A^2 \rangle \geq \langle A \rangle^2.$

(1.10)

An interesting geometrical interpretation of the uncertainty goes as follows. Consider the one-dimensional vector subspace U_Ψ generated by Ψ . Take the state $A\Psi$ and project it to the subspace U_Ψ . The projection, we claim is $\langle A \rangle\Psi$ and the part of $A\Psi$ in the orthogonal subspace U_Ψ^\perp is a vector of norm equal to the uncertainty ΔA . Indeed the orthogonal projector $P_{U_\Psi^\perp}$ is

$$P_{U_\Psi^\perp} = |\Psi\rangle\langle\Psi|, \quad (1.11)$$

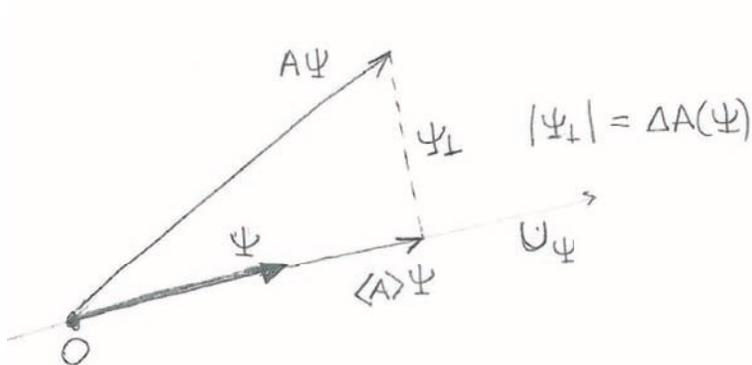


Figure 1: A state Ψ and the one-dimensional subspace U_Ψ generated by it. The projection of $A\Psi$ to U_Ψ is $\langle A \rangle \Psi$. The orthogonal complement Ψ_\perp is a vector whose norm is the uncertainty $\Delta A(\Psi)$.

so that

$$P_{U_\Psi} A|\Psi\rangle = |\Psi\rangle\langle\Psi|A|\Psi\rangle = |\Psi\rangle\langle A\rangle. \quad (1.12)$$

Moreover, the vector $A|\Psi\rangle$ minus its projection must be a vector $|\Psi_\perp\rangle$ orthogonal to $|\Psi\rangle$

$$A|\Psi\rangle - \langle A \rangle|\Psi\rangle = |\Psi_\perp\rangle, \quad (1.13)$$

as is easily confirmed by taking the overlap with the bra Ψ . Since the norm of the above left-hand side is the uncertainty, we confirm that $\Delta A = |\Psi_\perp|$, as claimed. These results are illustrated in Figure 1.

2 The Uncertainty Principle

The uncertainty principle is an inequality that is satisfied by the product of the uncertainties of two Hermitian operators that fail to commute. Since the uncertainty of an operator on any given physical state is a number greater than or equal to zero, the product of uncertainties is also a real number greater than or equal to zero. The uncertainty inequality often gives us a lower bound for this product.

When the two operators in question commute, the uncertainty inequality gives no information.

Let us state the uncertainty inequality. Consider two Hermitian operators A and B and a physical state Ψ of the quantum system. Let ΔA and ΔB denote the uncertainties of A and B , respectively, in the state Ψ . Then we have

$$\boxed{(\Delta A)^2(\Delta B)^2 \geq \left(\langle \Psi | \frac{1}{2i} [A, B] | \Psi \rangle \right)^2.} \quad (2.14)$$

The left hand side is a real, non-negative number. For this to be consistent inequality, the right-hand side must also be a real number that is not negative. Since the right-hand side appears squared, the object inside the parenthesis must be real. This can only happen for all Ψ if the operator

$$\frac{1}{2i} [A, B] \quad (2.15)$$

is Hermitian. For this first note that the commutator of two Hermitian operators is *anti*-Hermitian:

$$[A, B]^\dagger = (AB)^\dagger - (BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA = -[A, B] \quad (2.16)$$

The presence of the i then makes the operator in (2.15) Hermitian. Note that the uncertainty inequality can also be written as

$$\Delta A \Delta B \geq \left| \langle \Psi | \frac{1}{2i} [A, B] | \Psi \rangle \right|. \quad (2.17)$$

where the bars on the right-hand side denote absolute value.

Before we prove the theorem, let's do the canonical example! Substituting \hat{x} for A and \hat{p} for B results in the position-momentum uncertainty relation you have certainly worked with:

$$(\Delta x)^2 (\Delta p)^2 \geq \left(\langle \Psi | \frac{1}{2i} [\hat{x}, \hat{p}] | \Psi \rangle \right)^2. \quad (2.18)$$

Since $[\hat{x}, \hat{p}]/(2i) = \hbar/2$ we get

$$(\Delta x)^2 (\Delta p)^2 \geq \frac{\hbar^2}{4} \quad \rightarrow \quad \Delta x \Delta p \geq \frac{\hbar}{2}. \quad (2.19)$$

We are interested in the proof of the uncertainty inequality for it gives the information that is needed to find the conditions that lead to saturation.

Proof. We define the following two states:

$$\begin{aligned} |f\rangle &\equiv (A - \langle A \rangle I) |\Psi\rangle \\ |g\rangle &\equiv (B - \langle B \rangle I) |\Psi\rangle. \end{aligned} \quad (2.20)$$

Note that by the definition (1.2) of uncertainty,

$$\begin{aligned} \langle f|f\rangle &= (\Delta A)^2, \\ \langle g|g\rangle &= (\Delta B)^2. \end{aligned} \quad (2.21)$$

The Schwarz inequality immediately furnishes us an inequality involving precisely the uncertainties

$$\langle f|f\rangle \langle g|g\rangle \geq |\langle f|g\rangle|^2, \quad (2.22)$$

and therefore we have

$$(\Delta A)^2 (\Delta B)^2 \geq |\langle f|g\rangle|^2 = (\text{Re}\langle f|g\rangle)^2 + (\text{Im}\langle f|g\rangle)^2. \quad (2.23)$$

Writing $\check{A} = (A - \langle A \rangle I)$ and $\check{B} = (B - \langle B \rangle I)$, we now begin to compute the right-hand side:

$$\langle f|g\rangle = \langle \Psi | \check{A} \check{B} | \Psi \rangle = \langle \Psi | (A - \langle A \rangle I)(B - \langle B \rangle I) | \Psi \rangle = \langle \Psi | AB | \Psi \rangle - \langle A \rangle \langle B \rangle, \quad (2.24)$$

and since $|f\rangle$ and $|g\rangle$ go into each other as we exchange A and B ,

$$\langle g|f\rangle = \langle \Psi | \check{A} \check{B} | \Psi \rangle = \langle \Psi | BA | \Psi \rangle - \langle B \rangle \langle A \rangle. \quad (2.25)$$

From the two equations above we find a nice expression for the imaginary part of $\langle f|g\rangle$:

$$\text{Im}\langle f|g\rangle = \frac{1}{2i}(\langle f|g\rangle - \langle g|f\rangle) = \frac{1}{2i}\langle\Psi|[A, B]|\Psi\rangle. \quad (2.26)$$

For the real part the expression is not that simple, so it is best to leave it as the anticommutator of the checked operators:

$$\text{Re}\langle f|g\rangle = \frac{1}{2}(\langle f|g\rangle + \langle g|f\rangle) = \frac{1}{2}\langle\Psi|\{\hat{A}, \hat{B}\}|\Psi\rangle \quad (2.27)$$

Back in (2.23) we get

$$(\Delta A)^2(\Delta B)^2 \geq \left(\langle\Psi|\frac{1}{2i}[A, B]|\Psi\rangle\right)^2 + \left(\langle\Psi|\frac{1}{2}\{\hat{A}, \hat{B}\}|\Psi\rangle\right)^2. \quad (2.28)$$

This can be viewed as the most complete form of the uncertainty inequality. It turns out, however, that the second term on the right hand side is seldom simple enough to be of use, and many times it can be made equal to zero for certain states. At any rate, the term is positive or zero so it can be dropped while preserving the inequality. This is often done, thus giving the celebrated form (2.14) that we have now established.

Now that we have proven the uncertainty inequality, we can ask: What are the conditions for this inequality to be saturated? If the goal is to minimize uncertainties, under what conditions can we achieve the minimum possible product of uncertainties? As the proof shows, saturation is achieved under two conditions:

1. The Schwarz inequality is saturated. For this we need $|g\rangle = \beta|f\rangle$ where $\beta \in \mathbb{C}$.
2. $\text{Re}(\langle f|g\rangle) = 0$, so that the last term in (2.28) vanishes. This means that $\langle f|g\rangle + \langle g|f\rangle = 0$.

Using $|g\rangle = \beta|f\rangle$ in Condition 2, we get

$$\langle f|g\rangle + \langle g|f\rangle = \beta\langle f|f\rangle + \beta^*\langle f|f\rangle = (\beta + \beta^*)\langle f|f\rangle = 0, \quad (2.29)$$

which requires $\beta + \beta^* = 0$ or that the real part of β vanish. It follows that β must be purely imaginary. So, $\beta = i\lambda$, with λ real, and therefore the uncertainty inequality will be saturated if and only if

$$|g\rangle = i\lambda|f\rangle, \quad \lambda \in \mathbb{R}. \quad (2.30)$$

More explicitly this requires

Saturation Condition: $(B - \langle B\rangle I)|\Psi\rangle = i\lambda(A - \langle A\rangle I)|\Psi\rangle.$

(2.31)

This must be viewed as a condition for Ψ , given any two operators A and B . Moreover, note that $\langle A\rangle$ and $\langle B\rangle$ are Ψ dependent. What is λ , physically? Well, the norm of λ is actually fixed by the equation. Taking the norm of both sides we get

$$\Delta B = |\lambda|\Delta A \quad \rightarrow \quad |\lambda| = \frac{\Delta B}{\Delta A}. \quad (2.32)$$

The classic illustration of this saturation condition is worked out for the x, p uncertainty inequality $\Delta x\Delta p \geq \hbar/2$. You will find that gaussian wavefunctions satisfy the saturation condition.

3 The Energy-Time uncertainty

A more subtle form of the uncertainty relation deals with energy and time. The inequality is sometimes stated vaguely in the form $\Delta E \Delta t \gtrsim \hbar$. In here there is no problem in defining ΔE precisely, after all we have the Hamiltonian operator, and its uncertainty ΔH is a perfect candidate for the ‘energy uncertainty’. The problem is time. Time is not an operator in quantum mechanics, it is a parameter, a real number used to describe the way systems change. Unless we define Δt in a precise way we cannot hope for a well-defined uncertainty relation.

We can try a rough, heuristic definition, in order to illustrate the spirit of the inequality. Consider a photon that is detected at some point in space, as a passing oscillatory wave of exact duration T . Even without quantum mechanical considerations we can ask the observer what was the angular frequency ω of the pulse. In order to answer our question the observer will attempt to count the number N of complete oscillations of the waveform that went through. Of course, this number N is given by T divided by the period $2\pi/\omega$ of the wave:

$$N = \frac{\omega T}{2\pi}. \quad (3.33)$$

The observer, however, will typically fail to count full waves, because as the pulse gets started from zero and later on dies off completely, the waveform will cease to follow the sinusoidal pattern. Thus we expect an uncertainty $\Delta N \gtrsim 1$. Given the above relation, this implies an uncertainty $\Delta\omega$ in the value of the angular frequency

$$\Delta\omega T \gtrsim 2\pi. \quad (3.34)$$

This is all still classical, the above identity is something electrical engineers are well aware of. It represents a limit on the ability to ascertain accurately the frequency of a wave that is observed for a limited amount of time. This becomes quantum mechanical if we speak of a single photon, whose energy is $E = \hbar\omega$. Then $\Delta E = \hbar\Delta\omega$, so that multiplying the above inequality by \hbar we get

$$\Delta E T \gtrsim h. \quad (3.35)$$

In this uncertainty inequality T is the duration of the pulse. It is a reasonable relation but the presence of \gtrsim betrays its lack of full precision.

We can find a precise energy/ Q -ness uncertainty inequality by applying the general uncertainty inequality to the Hamiltonian H and another Hermitian operator Q , as did the distinguished Russian physicists L. Mandelstam and Tamm shortly after the formulation of the uncertainty principle. We would then have

$$\Delta H \Delta Q \geq \left| \langle \Psi | \frac{1}{2i} [H, Q] | \Psi \rangle \right|. \quad (3.36)$$

This starting point is interesting because the commutator $[H, Q]$ encodes something very physical about Q . Indeed, let us consider henceforth the case in which the operator Q has *no time dependence*. It could be, for example some function of \hat{x} and \hat{p} , or for a spin-1/2 particle, the operator $|+\rangle\langle-|$. Such

operator Q can easily have time-dependent expectation values, but the time dependence originates from the time dependence of the states, not from the operator Q itself.

To explore the meaning of $[H, Q]$ we begin by computing the time-derivative of the expectation value of Q :

$$\frac{d}{dt}\langle Q \rangle = \frac{d}{dt}\langle \Psi, Q\Psi \rangle = \left\langle \frac{\partial \Psi}{\partial t}, Q\Psi \right\rangle + \left\langle \Psi, Q \frac{\partial \Psi}{\partial t} \right\rangle \quad (3.37)$$

where we did not have to differentiate Q as it is time-independent. At this point we can use the Schrödinger equation to find

$$\begin{aligned} \frac{d}{dt}\langle Q \rangle &= \left\langle \frac{1}{i\hbar} H\Psi, Q\Psi \right\rangle + \left\langle \Psi, Q \frac{1}{i\hbar} H\Psi \right\rangle \\ &= \frac{i}{\hbar} \left(\langle H\Psi, Q\Psi \rangle - \langle \Psi, QH\Psi \rangle \right) \\ &= \frac{i}{\hbar} \langle \Psi, (HQ - QH)\Psi \rangle = \frac{i}{\hbar} \langle \Psi, [H, Q]\Psi \rangle \end{aligned} \quad (3.38)$$

where we used the Hermiticity of the Hamiltonian. We have thus arrived at

$$\frac{d}{dt}\langle Q \rangle = \frac{i}{\hbar} \langle [H, Q] \rangle \quad \text{for time-independent } Q. \quad (3.39)$$

This is a very important result. Each time you see $[H, Q]$ you should think ‘time derivative of $\langle Q \rangle$ ’. In classical mechanics one usually looks for conserved quantities, that is, functions of the dynamical variables that are time independent. In quantum mechanics a conserved operator is one whose expectation value is time independent. An operator Q is conserved if it commutes with the Hamiltonian!

With this result, the inequality (3.36) can be simplified. Indeed, using (3.39) we have

$$\left| \left\langle \frac{1}{2i} [H, Q] \right\rangle \right| = \left| \frac{1}{2i} \frac{\hbar}{i} \frac{d\langle Q \rangle}{dt} \right| = \frac{\hbar}{2} \left| \frac{d\langle Q \rangle}{dt} \right| \quad (3.40)$$

and therefore

$$\Delta H \Delta Q \geq \frac{\hbar}{2} \left| \frac{d\langle Q \rangle}{dt} \right|, \quad \text{for time-independent } Q. \quad (3.41)$$

This is a perfectly precise uncertainty inequality. The terms in it suggest a definition of a time Δt_Q

$$\Delta t_Q \equiv \frac{\Delta Q}{\left| \frac{d\langle Q \rangle}{dt} \right|}. \quad (3.42)$$

This quantity has units of time. It is the time it would take $\langle Q \rangle$ to change by ΔQ if both ΔQ and the velocity $\frac{d\langle Q \rangle}{dt}$ were time-independent. Since they are not necessarily so, we can view Δt_Q as the time for ‘‘appreciable’’ change in $\langle Q \rangle$. This is certainly so when $\langle Q \rangle$ and ΔQ are roughly of the same size. In terms of Δt_Q the uncertainty inequality reads

$$\Delta H \Delta t_Q \geq \frac{\hbar}{2}. \quad (3.43)$$

This is still a precise inequality, given that Δt_Q has a concrete definition in (3.42).

As you will consider in the homework, (3.41) can be used to derive an inequality for time Δt_\perp that it takes for a system to become orthogonal to itself. If we call the initial state $\Psi(0)$, we call Δt_\perp the smallest time for which $\langle \Psi(0), \Psi(\Delta t_\perp) \rangle = 0$. You will be able to show that

$$\Delta H \Delta t_\perp \geq \frac{\hbar}{4}. \quad (3.44)$$

The speed in which a state can turn orthogonal depends on the energy uncertainty, and in quantum computation it plays a role in limiting the maximum possible speed of a computer for a fixed finite energy.

The uncertainty relation involves ΔH . It is natural to ask if this quantity is time dependent. As we show now, it is not, if the Hamiltonian is a time-independent operator. Indeed, if H is time independent, we can use H and H^2 for Q in (3.39) so that

$$\begin{aligned} \frac{d}{dt} \langle H \rangle &= \frac{i}{\hbar} \langle [H, H] \rangle = 0, \\ \frac{d}{dt} \langle H^2 \rangle &= \frac{i}{\hbar} \langle [H, H^2] \rangle = 0. \end{aligned} \quad (3.45)$$

It then follows that

$$\frac{d}{dt} (\Delta H)^2 = \frac{d}{dt} (\langle H^2 \rangle - \langle H \rangle^2) = 0. \quad (3.46)$$

showing that ΔH is a constant. So we have shown that

If H is time independent, the uncertainty ΔH is constant in time.

(3.47)

The concept of conservation of energy uncertainty can be used to understand some aspects of atomic decays. Consider, for illustration the hyperfine transition in the hydrogen atom. Due to the existence of proton spin and the electron spin, the ground state of hydrogen is fourfold degenerate, corresponding to the four possible combinations of spins (up-up, up-down, down-up, down-down). The magnetic interaction between the spins actually breaks this degeneracy and produces the so-called “hyperfine” splitting. This is a very tiny split: $\delta E = 5.88 \times 10^{-6} \text{ eV}$ (compare with about 13.6 eV for the ground state energy). For a hyperfine atomic transition, the emitted photon carries the energy difference δE resulting in a wavelength of 21.1cm and a frequency $\nu = 1420.405751786(30) \text{ MHz}$. The eleven significant digits of this frequency attest to the sharpness of the emission line. The issue of uncertainty arises because the excited state of the hyperfine splitting has a lifetime τ_H for decay to the ground state and emission of a photon. This lifetime is extremely long, in fact $\tau_H \sim 11$ million years ($= 3.4 \times 10^{14}$ sec, recalling that a year is about $\pi \times 10^7$ sec, accurate to better than 1%). This lifetime can be viewed as the time that takes some observable of the electron-proton system to change significantly (its total spin angular momentum, perhaps) so by the uncertainty principle it must be related to some energy uncertainty $\Delta E \sim \hbar/\tau_H \simeq 2 \times 10^{-30} \text{ eV}$. of the original excited state of the

hydrogen atom. Once the decay takes place the atom goes to the fully stable ground state, without any possible energy uncertainty. By the conservation of energy uncertainty, the photon must carry the uncertainty ΔE . But $\Delta E/\delta E \sim 3 \times 10^{-25}$, an absolutely infinitesimal effect on the photon. There is no broadening of the 21 cm line! That's one reason it is so useful in astronomy. For decays with much shorter lifetimes there can be an observable broadening of an emission line due to the energy-time uncertainty principle.

4 Lower bounds for ground state energies

You may recall that the variational principle could be used to find *upper* bounds on ground state energies. The uncertainty principle can be used to find *lower* bounds for the ground state energy of certain systems. Use below the uncertainty principle in the form $\Delta x \Delta p \geq \hbar/2$ to find rigorous lower bounds for the ground state energy of one-dimensional Hamiltonians. This is best illustrated by an example.

Consider a particle in a one-dimensional quartic potential considered earlier

$$H = \frac{p^2}{2m} + \alpha x^4, \quad (4.48)$$

where $\alpha > 0$ is a constant with units of energy over length to the fourth power. Our goal is to find a *lower bound* for the ground state energy $\langle H \rangle_{gs}$. Taking the ground state expectation value of the Hamiltonian we have

$$\langle H \rangle_{gs} = \frac{\langle p^2 \rangle_{gs}}{2m} + \alpha \langle x^4 \rangle_{gs}, \quad (4.49)$$

Recalling that

$$(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2, \quad (4.50)$$

we see that

$$\langle p^2 \rangle \geq (\Delta p)^2, \quad (4.51)$$

for any state of the system. We should note however, that for the ground state (or any bound state) $\langle p \rangle = 0$ so that in fact

$$\langle p^2 \rangle_{gs} = (\Delta p)_{gs}^2, \quad (4.52)$$

From the inequality $\langle A^2 \rangle \geq \langle A \rangle^2$ we have

$$\langle x^4 \rangle \geq \langle x^2 \rangle^2. \quad (4.53)$$

Moreover, just like for momentum above, $(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2$ leads to

$$\langle x^2 \rangle \geq (\Delta x)^2, \quad (4.54)$$

so that

$$\langle x^4 \rangle \geq (\Delta x)^4, \quad (4.55)$$

for the expectation value on arbitrary states. Therefore

$$\langle H \rangle_{gs} = \frac{\langle p^2 \rangle_{gs}}{2m} + \alpha \langle x^4 \rangle_{gs} \geq \frac{(\Delta p_{gs})^2}{2m} + \alpha (\Delta x_{gs})^4 \quad (4.56)$$

From the uncertainty principle

$$\Delta x_{gs} \Delta p_{gs} \geq \frac{\hbar}{2} \quad \rightarrow \quad \Delta p_{gs} \geq \frac{\hbar}{2\Delta x_{gs}}. \quad (4.57)$$

Back to the value of $\langle H \rangle_{gs}$ we get

$$\langle H \rangle_{gs} \geq \frac{\hbar^2}{8m(\Delta x_{gs})^2} + \alpha (\Delta x_{gs})^4. \quad (4.58)$$

The quantity to the right of the inequality is a function of Δx_{gs} . This function has been plotted in Figure 2.

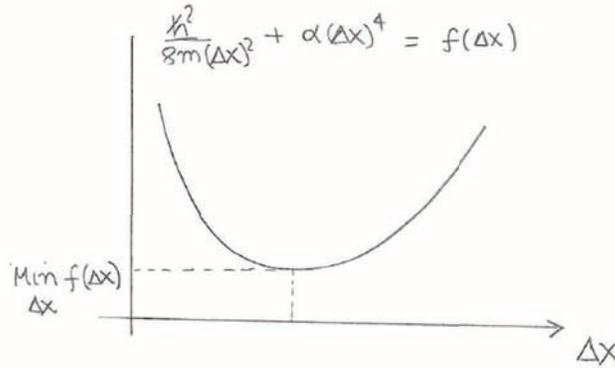


Figure 2: We have that $\langle H_{gs} \rangle \geq f(\Delta x_{gs})$ but we don't know the value of Δx_{gs} . As a result, we can only be certain that $\langle H_{gs} \rangle$ is greater than or equal to the *lowest* value the function $f(\Delta x_{gs})$ can take.

If we knew the value of Δx_{gs} we would immediately know that $\langle H \rangle_{gs}$ is bigger than the value taken by the right-hand side. This would be quite nice, since we want the highest possible lower bound. Since we don't know the value of Δx_{gs} , however, the only thing we can be sure of is that $\langle H \rangle_{gs}$ is bigger than the *lowest* value that can be taken by the expression to the right of the inequality as we vary Δx_{gs} :

$$\langle H \rangle_{gs} \geq \text{Min}_{\Delta x} \left(\frac{\hbar^2}{8m(\Delta x)^2} + \alpha (\Delta x)^4 \right). \quad (4.59)$$

The minimization problem is straightforward. In fact

$$f(x) = \frac{A}{x^2} + Bx^4 \text{ is minimized for } x^2 = 2^{-1/3} \left(\frac{A}{B} \right)^{1/3} \text{ yielding } f = 2^{1/3} \frac{3}{2} (A^2 B)^{1/3}. \quad (4.60)$$

Applied to (4.59) we obtain

$$\langle H \rangle_{gs} \geq 2^{1/3} \frac{3}{8} \left(\frac{\hbar^2 \sqrt{\alpha}}{m} \right)^{2/3} \simeq 0.4724 \left(\frac{\hbar^2 \sqrt{\alpha}}{m} \right)^{2/3}. \quad (4.61)$$

This is the final lower bound for the ground state energy. It is actually not too bad, for the ground state instead of the prefactor 0.4724, we have 0.668.

5 Diagonalization of Operators

When we have operators we wish to understand, it can be useful to find a basis on the vector space for which the operators are represented by matrices that take a simple form. Diagonal matrices are matrices where all non diagonal entries vanish. If we can find a set of basis vectors for which the matrix representing an operator is diagonal we say that the operator is **diagonalizable**.

If an operator T is diagonal in some basis (u_1, \dots, u_n) of the vector space V , its matrix takes the form $\text{diag}(\lambda_1, \dots, \lambda_n)$, with constants λ_i , and we have

$$Tu_1 = \lambda_1 u_1, \quad \dots, \quad Tu_n = \lambda_n u_n. \quad (5.62)$$

The basis vectors are recognized as eigenvectors with eigenvalues given by the diagonal elements. It follows that *a matrix is diagonalizable if and only if it possesses a set of eigenvectors that span the vector space*. Recall that all operators T on complex vector spaces have at least one eigenvalue and thus at least a one eigenvector. But not even in complex vector spaces all operators have enough eigenvectors to span the space. Those operators cannot be diagonalized. The simplest example of such operator is provided by the two-by-two matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \quad (5.63)$$

The only eigenvalue of this matrix is $\lambda = 0$ and the associated eigenvector is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Since a two-dimensional vector space cannot be spanned with one eigenvector, this matrix cannot be diagonalized. Having seen that the question of diagonalization of an operator is ultimately a question about its eigenvectors, we want to emphasize that the question can be formulated without referring to any basis. Bases, of course are useful, to express concretely

Suppose we have a vector space V and we have chosen a basis (v_1, \dots, v_n) such that a linear operator has a matrix representation $T_{ij}(\{v\})$ that is not diagonal. As we learned before, if we change basis to a new one (u_1, \dots, u_n) using a linear operator A such that

$$u_k = A v_k, \quad (5.64)$$

the matrix representation $T_{ij}(\{u\})$ of the operator in the new basis takes the form

$$T(\{u\}) = A^{-1}T(\{v\})A \quad \text{or} \quad T_{ij}(\{u\}) = (A^{-1})_{ik}T_{kp}(\{v\})A_{pj}, \quad (5.65)$$

where the matrix A_{ij} is the representation of A in the original v -basis. The operator T is diagonalizable if there is an operator A such that $T_{ij}(\{u\})$ is diagonal.

There are two pictures of the diagonalization: One can consider the operator T and state that its matrix representation is diagonal when referred to the u basis obtained by acting with A on the original v basis. Alternatively, we can view the result as the existence of a related operator $A^{-1}TA$ that is diagonal in the *original* v basis. Indeed, $Tu_i = \lambda_i u_i$ (i not summed) implies that $TA v_i = \lambda_i A v_i$

and acting with A^{-1} that $(A^{-1}TA)v_i = \lambda_i v_i$, which confirms that $A^{-1}TA$ is represented by a diagonal matrix in the original v basis. Both viewpoints are valuable.

It is useful to note that the columns of the matrix A are in fact the eigenvectors of $T(\{v\})$. We see this as follows. Since the eigenvectors are the u_k we have

$$u_k = Av_k \quad \rightarrow \quad u_k = \sum_i A_{ik}v_i. \quad (5.66)$$

Using the original basis means v_i is represented by a column vector of zeroes with a single unit entry at the i -th position. We thus find

$$u_k = \begin{pmatrix} A_{1k} \\ \vdots \\ A_{nk} \end{pmatrix}. \quad (5.67)$$

confirming that the k -th column of A is the k -th eigenvector of T .

While not all operators on complex vector spaces can be diagonalized, the situation is much improved for Hermitian operators. Recall that T is Hermitian if $T = T^\dagger$. Hermitian operators can be diagonalized, and so can unitary operators. But even more is true: the operators take diagonal form in an orthonormal basis!

An operator M is said to be **unitarily** diagonalizable if there is an *orthonormal* basis in which its matrix representation is a diagonal matrix. That basis, therefore, is an orthonormal basis of eigenvectors. Starting with an arbitrary orthonormal basis (e_1, \dots, e_n) where the matrix representation of M is $M(\{e\})$, a unitary transformation of this basis produces the orthonormal basis in which the operator takes diagonal form. More explicitly, there is a unitary matrix U ($U^\dagger = U^{-1}$) and a diagonal matrix D_M such that

$$U^\dagger M(\{e\}) U = D_M. \quad (5.68)$$

6 The Spectral Theorem

While we could prove, as most textbooks do, that Hermitian operators are unitarily diagonalizable, this result holds for a more general class of operators, called normal operators. The proof is not harder than the one for hermitian operators. An operator M is said to be **normal** if it commutes with its adjoint:

$$M \text{ is normal : } [M^\dagger, M] = 0. \quad (6.69)$$

Hermitian operators are clearly normal. So are anti-hermitian operators ($M^\dagger = -M$ is antihermitian). Unitary operators U are normal because both $U^\dagger U$ and UU^\dagger are equal to the identity matrix and thus U and U^\dagger commute.

Exercise. If an operator M is normal show that so is $V^\dagger M V$ where V is a unitary operator.

Lemma: Let w be an eigenvector of the normal operator M : $Mw = \lambda w$. Then w is also an eigenvector of M^\dagger with complex conjugate eigenvalue:

$$M^\dagger w = \lambda^* w. \quad (6.70)$$

Proof: Define $u = (M^\dagger - \lambda^* I)w$. The result holds if u is the zero vector. To show this we compute the norm-squared of u :

$$|u|^2 = \langle u, u \rangle = \langle (M^\dagger - \lambda^* I)w, (M^\dagger - \lambda^* I)w \rangle \quad (6.71)$$

Using the adjoint property to move the operator in the first entry to the second entry:

$$|u|^2 = \langle w, (M - \lambda I)(M^\dagger - \lambda^* I)w \rangle \quad (6.72)$$

Since M and M^\dagger commute, so do the two factors in parenthesis and therefore

$$|u|^2 = \langle w, (M^\dagger - \lambda^* I)(M - \lambda I)w \rangle = 0 \quad (6.73)$$

since $(M - \lambda I)$ kills w . It follows that $u = 0$ and therefore (6.70) holds. \square

We can now state our main theorem, called the *spectral theorem*. It states that a matrix is unitarily diagonalizable if and only if it is normal. More to the point,

Spectral Theorem: Let M be an operator in a complex vector space. The vector space has a orthonormal basis comprised of eigenvectors of M if and only if M is normal.

(6.74)

Proof. It is easy to show that unitarily diagonalizable implies normality. Indeed, from (5.68) and dropping the reference to the e -basis,

$$M = UD_M U^\dagger \quad \text{and therefore} \quad M^\dagger = UD_M^\dagger U^\dagger.$$

We then get

$$M^\dagger M = UD_M^\dagger D_M U^\dagger \quad \text{and} \quad MM^\dagger = UD_M D_M^\dagger U^\dagger.$$

so that

$$[M^\dagger, M] = U(D_M^\dagger D_M - D_M D_M^\dagger)U^\dagger = 0,$$

because any two diagonal matrices commute.

Now let us prove that M provides a basis of orthonormal eigenvectors. The proof is by induction. The result is clearly true for $\dim V = 1$. We assume that it holds for $(n-1)$ -dimensional vector spaces and consider the case of n -dimensional V . Let M be an $n \times n$ matrix referred to the orthonormal basis $(|1\rangle, \dots, |n\rangle)$ of V so that $M_{ij} = \langle i|M|j\rangle$. We know there is at least one eigenvalue λ_1 with a non-zero eigenvector $|x_1\rangle$ of unit norm:

$$M|x_1\rangle = \lambda_1|x_1\rangle \quad \text{and} \quad M^\dagger|x_1\rangle = \lambda_1^*|x_1\rangle, \quad (6.75)$$

in view of the Lemma. There is, we claim, a unitary matrix U_1 such that

$$|x_1\rangle = U_1|1\rangle \quad \rightarrow \quad U_1^\dagger|x_1\rangle = |1\rangle. \quad (6.76)$$

U_1 is not unique and can be constructed as follows: extend $|x_1\rangle$ to an orthonormal basis $|x_1\rangle, \dots, |x_N\rangle$ using Gram-Schmidt. Then write $U_1 = \sum_i |x_i\rangle\langle i|$. Define now

$$M_1 \equiv U_1^\dagger M U_1. \quad (6.77)$$

M_1 is also normal and $M_1|1\rangle = U_1^\dagger M U_1|1\rangle = U_1^\dagger M|x_1\rangle = \lambda_1 U_1^\dagger|x_1\rangle = \lambda_1|1\rangle$, so that

$$M_1|1\rangle = \lambda_1|1\rangle. \quad (6.78)$$

Let us now examine the explicit form of the matrix M_1 :

$$\langle j|M_1|1\rangle = \lambda_1\langle j|1\rangle = \lambda_1\delta_{i,j}, \quad (6.79)$$

which says that the first column of M_1 has zeroes in all entries except the first. Moreover

$$\langle 1|M_1|j\rangle = (\langle j|M_1^\dagger|1\rangle)^* = (\lambda_1^*\langle j|1\rangle)^* = \lambda_1\langle 1|j\rangle = \lambda_1\delta_{i,j}, \quad (6.80)$$

where we used $M_1^\dagger|1\rangle = \lambda_1^*|1\rangle$ which follows from the normality of M_1 . It follows from the two last equations that M_1 , in the original basis, takes the form

$$M_1 = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & M' & \\ 0 & & & \end{array} \right).$$

Since M_1 is normal, one can see that M' is a normal $(n-1)$ -by- $(n-1)$ matrix. By the induction hypothesis M' can be unitarily diagonalized so that $U'^\dagger M' U'$ is diagonal for some $(n-1)$ -by- $(n-1)$ unitary matrix U' . The matrix U' can be extended to an n -by- n unitary matrix \hat{U} as follows

$$\hat{U} = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{array} \right). \quad (6.81)$$

It follows that $\hat{U}^\dagger M_1 \hat{U} = \hat{U}^\dagger U_1^\dagger M U_1 \hat{U} = (U_1 \hat{U})^\dagger M (U_1 \hat{U})$ is diagonal, proving the desired result. \square .

Of course this theorem implies that Hermitian and unitary operators are unitarily diagonalizable. In other words the eigenvectors form an orthonormal basis. This is true whether or not there are degeneracies in the spectrum. The proof does not require discussion of this as a special case. If an eigenvalue of M is degenerate and appears k times, then there are k orthonormal eigenvectors associated with the corresponding k -dimensional M -invariant subspace of the vector space.

We conclude this section with a description of the general situation that we may encounter when diagonalizing a normal operator T . In general, we expect degeneracies in the eigenvalues so that each eigenvalue λ_k is repeated $d_k \geq 1$ times. An eigenvalue λ_k is degenerate if $d_k > 1$. It follows that V has T -invariant subspaces of different dimensionalities. Let U_k denote the T -invariant subspace of dimension $d_k \geq 1$ spanned by eigenvectors with eigenvalue λ_k :

$$U_k \equiv \{v \in V \mid T v = \lambda_k v\}, \quad \dim U_k = d_k. \quad (6.82)$$

By the spectral theorem U_k has a basis comprised by d_k orthonormal eigenvectors $(u_1^{(k)}, \dots, u_{d_k}^{(k)})$. Note that while the addition of eigenvectors with different eigenvalues does not give eigenvectors, in the subspace U_k all vectors are eigenvectors with the same eigenvalue, and that's why addition makes sense, U_k as defined is a vector space, and adding eigenvectors in U_k gives eigenvectors. The full space V is decomposed as the direct sum of the invariant subspaces of T :

$$V = U_1 \oplus U_2 \oplus \dots \oplus U_m, \quad \dim V = \sum_{i=1}^m d_i, \quad m \geq 1. \quad (6.83)$$

All U_i subspaces are guaranteed to be orthogonal to each other. In fact the full list of eigenvectors is a list of orthonormal vectors that form a basis for V is conveniently ordered as follows:

$$(u_1^{(1)}, \dots, u_{d_1}^{(1)}, \dots, u_1^{(m)}, \dots, u_{d_m}^{(m)}). \quad (6.84)$$

The matrix T is manifestly diagonal in this basis because each vector above is an eigenvector of T and is orthogonal to all others. The matrix representation of T reads

$$T = \text{diag} \left(\underbrace{\lambda_1, \dots, \lambda_1}_{d_1 \text{ times}}, \dots, \underbrace{\lambda_m, \dots, \lambda_m}_{d_m \text{ times}} \right) \quad (6.85)$$

This is clear because the first d_1 vectors in the list are in U_1 , the second d_2 vectors are in U_2 , and so on and so forth until the last d_m vectors are in U_m .

If we had no degeneracies in the spectrum the basis (6.84) (with $d_i = 1$ for all i) would be rather unique if we require the matrix representation of T to be unchanged. Each vector could be multiplied by a phase. On the other hand, with degeneracies that the list (6.84) can be changed considerably without changing the matrix representation of T . Let V_k be a unitary operator on U_k , for each $k = 1, \dots, m$. We claim that the following basis of eigenvectors leads to the same matrix T :

$$(V_1 u_1^{(1)}, \dots, V_1 u_{d_1}^{(1)}, \dots, V_m u_1^{(m)}, \dots, V_m u_{d_m}^{(m)}). \quad (6.86)$$

Indeed, this is still a collection of eigenvectors of T with each of them orthogonal to the rest. Moreover, the first d_1 vectors are in U_1 , the second d_2 vectors are in U_2 and so on and so forth. More explicitly, for example, within U_k

$$\langle V_k u_i^{(k)}, T(V_k u_j^{(k)}) \rangle = \lambda_k \langle V_k u_i^{(k)}, V_k u_j^{(k)} \rangle = \lambda_k \langle u_i^{(k)}, u_j^{(k)} \rangle = \lambda_k \delta_{ij} \quad (6.87)$$

showing that in the U_k subspace the matrix for T is still diagonal with all entries equal to λ_k .

7 Simultaneous Diagonalization of Hermitian Operators

We say that two operators S and T in a vector space V operators can be **simultaneously diagonalized** if there is some basis of V in which both the matrix representation of S and the matrix representation of T are diagonal. It then follows that each vector in this basis is an eigenvector of S and an eigenvector of T .

A necessary condition for simultaneous diagonalization is that the operators S and T commute. Indeed, if they can be simultaneously diagonalized there is a basis where both are diagonal and they manifestly commute. If the operators don't commute, this is a basis-independent statement and therefore a simultaneous diagonal presentation cannot exist. Since arbitrary linear operators S and T on a complex vector space cannot be diagonalized, the vanishing of $[S, T]$ does not guarantee simultaneous diagonalization. But if the operators are Hermitian it does, as we show now.

Theorem. If S and T are commuting Hermitian operators they can be simultaneously diagonalized.

Proof. The main complication is that degeneracies in the spectrum require an some discussion. Either both operators have degeneracies or one has no degeneracies. Without loss of generality we can assume that there are two cases to consider

- (i) There is no degeneracy in the spectrum of T or,
- (ii) Both T and S have degeneracies in their spectrum.

Consider case (i) first. Since T is non-degenerate there is a basis (u_1, \dots, u_n) of eigenvectors of T with different eigenvalues

$$Tu_i = \lambda_i u_i, \quad i \text{ not summed}, \quad \lambda_i \neq \lambda_j \text{ for } i \neq j. \quad (7.88)$$

We now want to understand what kind of vector is Su_i . For this we act with T on it

$$T(Su_i) = S(Tu_i) = S(\lambda_i u_i) = \lambda_i (S u_i), \quad (7.89)$$

It follows that Su_i is also an eigenvector of T with eigenvalue λ_i , thus it must equal u_i , up to scale,

$$Su_i = \omega_i u_i, \quad (7.90)$$

showing that u_i is also an eigenvector of S , this time with eigenvalue ω_i . Thus any eigenvector of T is also an eigenvector of S , showing that these operators are simultaneously diagonalizable.

Now consider case (ii). Since T has degeneracies, as explained in the previous section, we have a decomposition of V in T -invariant subspaces U_k spanned by eigenvectors:

$$\begin{aligned}
 U_k &\equiv \{u \mid Tu = \lambda_k u\}, \quad \dim U_k = d_k \quad V = U_1 \oplus \dots \oplus U_m, \\
 \text{orthonormal basis for } V &: (u_1^{(1)}, \dots, u_{d_1}^{(1)}, \dots, u_1^{(m)}, \dots, u_{d_m}^{(m)}). \\
 T &= \text{diag} \left(\underbrace{\lambda_1, \dots, \lambda_1}_{d_1 \text{ times}}, \dots, \underbrace{\lambda_m, \dots, \lambda_m}_{d_m \text{ times}} \right) \text{ in this basis.}
 \end{aligned} \quad (7.91)$$

We also explained that the alternative orthonormal basis of V

$$(V_1 u_1^{(1)}, \dots, V_1 u_{d_1}^{(1)}, \dots, V_m u_1^{(m)}, \dots, V_m u_{d_m}^{(m)}). \quad (7.92)$$

leads to the same matrix for T when each V_k is a unitary operator on U_k .

We now claim that the U_k are also S -invariant subspaces! To show this let $u \in U_k$ and examine the vector Su . We have

$$T(Su) = S(Tu) = \lambda_k Su \quad \rightarrow \quad Su \in U_k. \quad (7.93)$$

We use the subspaces U_k and the basis (7.91) to organize the matrix representation of S in blocks. It follows that this matrix must have *block-diagonal* form since each subspace is S -invariant and orthogonal to all other subspaces. We cannot guarantee, however, that S is diagonal within each square block because $Su_i^{(k)} \in U_k$ but we have no reason to believe that $Su_i^{(k)}$ points along $u_i^{(k)}$.

Since S restricted to each S -invariant subspace U_k is hermitian we can find an orthonormal basis of U_k in which the matrix S is diagonal. This new basis is unitarily related to the original basis $(u_1^{(k)}, \dots, u_{d_k}^{(k)})$ and thus takes the form $(V_k u_1^{(k)}, \dots, V_k u_{d_k}^{(k)})$ with V_k a unitary operator in U_k . Note that the eigenvalues of S in this block need not be degenerate. Doing this for each block, we find a basis of the form (7.92) in which S is diagonal. But T is still diagonal in this new basis, so both S and T have been simultaneously diagonalized. \square

Remarks:

1. Note that the above proof gives an algorithmic way to produce the common list of eigenvectors. One diagonalizes one of the matrices and constructs the second matrix in the basis of eigenvectors of the first. These second matrix is block diagonal, where the blocks are organized by the degeneracies in the spectrum of the first matrix. One must then diagonalize within the blocks and is guaranteed that the new basis that works for the second matrix also works for the first.
2. If we had to simultaneously diagonalize three different commuting Hermitian operators S_1, S_2 and S_3 , all of which have degenerate spectra, we would proceed as follows. We diagonalize S_1 and fix a basis in which S_1 is diagonal. In this basis we must find that S_2 and S_3 have exactly the same block structure. The corresponding block matrices are simply the matrix representations of S_2 and S_3 in each of the invariant spaces U_k appearing in the diagonalization of S_1 . Since S_2 and S_3 commute, their restrictions to U_k commute. These restrictions can be diagonalized simultaneously, as guaranteed by our theorem which works for two matrices. The new basis in U_k that makes the restriction of S_2 and S_3 diagonal, will not disturb the diagonal form of S_1 in this block. This is repeated for each block, until we get a common basis of eigenvectors.
3. An inductive algorithm is clear. If we know how to simultaneously diagonalize n commuting Hermitian operators we can diagonalize $n + 1$ of them, call them S_1, \dots, S_{n+1} , as follows. We diagonalize S_1 and then consider the remaining n operators in the basis that makes S_1 diagonal.

We are guaranteed a common block structure for the n operators. The problem becomes one of simultaneous diagonalization of n commuting Hermitian block matrices, which is assumed known by the induction argument.

Corollary. If $\{S_1, \dots, S_n\}$ is a set of mutually commuting Hermitian operators they can all be simultaneously diagonalized.

8 Complete Set of Commuting Observables

We have discussed the problem of finding eigenvectors and eigenvalues of a Hermitian operator S . This hermitian operator is thought as a quantum mechanical observable. The eigenvectors of S are physical states of the system in which the observable S can be measured without uncertainty. The result of the measurement is the eigenvalue associated with the eigenvector.

If the Hermitian operator S has a non-degenerate spectrum, all eigenvalues are different and we have a rather nice situation in which each eigenvector can be uniquely distinguished by labeling it with the corresponding eigenvalue of S . The physical quantity associated with the observable can be used to distinguish the various eigenstates. Moreover, these eigenstates provide an orthonormal basis for the full vector space. In this case the operator S provides a “complete set of commuting observables” or a CSCO, in short. The set here has just one observable, the operator S .

The situation is more nontrivial if the Hermitian operator S exhibits degeneracies in its spectrum. This means that V has an S -invariant subspace of dimension $d > 1$, spanned by orthonormal eigenvectors (u_1, \dots, u_d) all of which have S eigenvalue λ . This time, the eigenvalue of S does not allow us to distinguish or to label uniquely the basis eigenstates of the invariant subspace. Physically this is a deficient situation, as we have explicitly different states – the various u_i 's – that we can't tell apart by the measurement of S alone. This time S does not provide a CSCO. Labeling eigenstates by the S eigenvalue does not suffice to distinguish them.

We are thus physically motivated to find another Hermitian operator T that is compatible with S . Two Hermitian operators are said to be **compatible observables** if they commute, since then we can find a basis of V comprised by simultaneous eigenvectors of the operators. These states can be labeled by two observables, namely, the two eigenvalues. If we are lucky, the basis eigenstates in each of the S -invariant subspaces of dimension higher than one can be organized into T eigenstates of different eigenvalues. In this case T breaks the spectral degeneracy of S and using T eigenvalues as well as S eigenvalues we can label uniquely a basis of orthonormal states of V . In this case we say that S and T form a CSCO.

We have now given enough motivation for a definition of a complete set of commuting observables. Consider a set of commuting observables, namely, a set $\{S_1, \dots, S_k\}$ of Hermitian operators acting on a complex vector space V that represents the physical state-space of some quantum system. By the theorem in the previous section, we can find an orthonormal basis of vectors in V such that each vector

is an eigenstate of every operator in the set. Assume that each eigenstate in the basis is labeled by the eigenvalues of the S_i operators. The set $\{S_1, \dots, S_k\}$ is said to be a **complete set of commuting observables** if no two states have the same labels.

It is a physically motivated assumption that for any physical quantum system there is a complete set of commuting observables, for otherwise there is no physical way to distinguish the various states that span the vector space. So in any physical problem we are urged to find such complete set, and we must include operators in such set until all degeneracies are broken. A CSCO need not be unique. Once we have a complete set of commuting observables, adding another observable causes no harm, although it is not necessary. Also, if (S_1, S_2) form a CSCO, so will $(S_1 + S_2, S_1 - S_2)$. Ideally, we want the smallest set of operators.

The first operator that is usually included in a CSCO is the Hamiltonian H . For bound state problems in one dimension, energy eigenstates are non-degenerate and thus the energy can be used to label uniquely the H -eigenstates. A simple example is the infinite square well. Another example is the one-dimensional harmonic oscillator. In such cases H forms the CSCO. If we have, however, a two-dimensional isotropic harmonic oscillator in the (x, y) plane, the Hamiltonian has degeneracies. At the first excited level we can have the first excited state of the x harmonic oscillator or, at the same energy, the first excited state of the y harmonic oscillator. We thus need another observable that can be used to distinguish these states. There are several options, as you will discuss in the homework.

MIT OpenCourseWare

<http://ocw.mit.edu>

8.05 Quantum Physics II

Fall 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.