Chapter 9. The Eigenvalue Problem for Bound Systems

Section 9.1. Some General Properties

§ 1 *Introduction and review.* We have proven in an earlier chapter several important results. The pure states of an observable O are eigenstates of the operator $\hat{O}$ that represents that observable; the values that the observable is allowed to take (the spectrum of the observable) are the eigenvalues of $\hat{O}$. In the early days of quantum mechanics many simple eigenvalue problems (i.e. particle in a box, the harmonic oscillator, the hydrogen atom) were solved by using Schrödinger version of quantum mechanics where the eigenvalue problem is a differential equation. For a while the Heisenberg-Jordan-Born version of quantum mechanics, expressed through infinite dimensional vectors and infinite dimensional matrices, was considered too cumbersome and non-intuitive. Very few people had the courage and the expertise to use it for solving specific problems. Nowadays we formulate the eigenvalue problem in the Schrödinger representation, pick a finite basis set, convert to an eigenvalue problem for a matrix, which we solve numerically. In this chapter we discuss how we solve eigenvalue problems for matrices.

All computer languages provide libraries with functions that given a matrix will give back its eigenvalues and eigenvectors. Because of this I will not discuss the methods used for finding matrix eigenvalues, except to explain

how this is done in `Mathematica`. Instead I will explain the basic theory and examine various aspects of the eigenvalue problems by using simple examples.

This chapter is confined to the eigenvalue problem for bound systems which have a discrete spectrum.

§ 2 *The eigenvalue problem for continuous spectrum is different.* Before we proceed I will explain the physical and mathematical differences between the eigenvalue problems for the discrete and that for the continuous spectrum. A system has a discrete spectrum when all particles are bound (meaning that they do not fly away from each other). This is certainly the case for the molecules studied in chemistry if their energy is less than the energy at which they break up into fragments. This statement needs to be qualified when dealing with *bound and finite but very large systems* which are often talked about as if they have a continuous spectrum. The simplest example that illustrates the issues involved is the energy of a particle in a very large box. There is no doubt that its spectrum is discrete, since we can calculate it exactly and it is

$$\frac{\hbar^2\pi^2n^2}{2mL^2} \text{ where } n = 1, 2, \ldots$$

However, when the box-size L is large, the difference between two successive energies is so small that we have no experiment with sufficient energy resolution to detect that the spectrum is discrete. Therefore in many experiments *it appears* that the spectrum is continuous. Nevertheless, we know better: this is not a continuous spectrum because we cannot change the energy by an infinitesimal amount (as long as L is finite). Sometimes people say that the spectrum of electrons in a piece of metal is continuous, but this is not strictly true either: if the piece of metal has finite size, each electron in it is enclosed

in a finite-size box and its spectrum is discrete; the spectrum only appears to be continuous. To understand the difference between a very dense spectrum (which appears to be continuous) and a truly continuous one, consider what happens if, in the same finite piece of metal, an electron absorbs a photon and has enough energy to escape from the metal. Now the electron can be located anywhere: in the metal or in the big, wide world outside. This system has "dissociated" into an electron and a positively charged piece of metal. The spectrum of this system is truly continuous because we can, at least in principle, change the energy of the photon by an infinitesimal amount and this will change the energy of the final state (electron plus ionized metal) by an infinitesimal amount. The same is true when a hydrogen atom is exposed to high-energy photons that ionize it: the energy spectrum of the ionized system is continuous. In general we encounter continuous spectra in systems that have so much energy that they break into fragments or when we study the collision of two systems (e.g. of two molecules). The pure states of systems that have a continuous spectrum are calculated by solving *the same eigenvalue problem* that we solve for calculating the discrete spectrum. For example, the pure energy states of the electron-proton pair formed by the ionization of the hydrogen atom are eigenstates of the energy operator. This is the same energy operator used to calculated the pure states of the atom (the electron bound to the proton). However, the *method of calculation* of the energy eigenstates of a bound system is very different from the method of calculation of the energy eigenstates of a dissociating (or colliding) system. The essential difference is in the *boundary conditions*. For a bound molecule the wave function must decay to zero at spatial points far from the molecule.

For a collision problem the wave function must be a linear superposition of two waves: one represents the molecules approaching each other, and the other the products of collision moving away from each other. For a stable molecule the wave function is localized in space. For an unbound system (having a continuous spectrum), the wave function is spread over the whole space. These differences in the physical situation translate into differences in the boundary conditions, which require a completely different method of solution of the eigenvalue problem. It is also significant that when we study collisions we control the total energy (we know it), which is conserved: the total energy after the collision is the same as the total energy prior to the collision. What we calculate is the probability that the particles produced by collision move in a certain direction, with a certain velocity, and have a certain internal energy (i.e. vibration, rotation). When we solve the energy eigenvalue problem for a bound state, the total energy of the system is not known, and it must be given by the calculation.

In this section we deal with the eigenvalue problem for *bound systems*. The method discussed here applies to bound molecules and also to a large piece of metal in spite of the fact that the spectrum of the latter appears to be continuous. As long as the electron stay inside the metal the eigenvalue problem is that of a bound system. Problems involving energies for which the system dissociates, or in which molecules collide, are treated when we discuss photochemistry or collision theory.

§ 3 *Some general properties of the eigenvalue problem for operators.* I repeat here some material that was discussed earlier, concerning those properties of the eigenvalue problem that can be derived from what we know

about operators. After that we move on and discuss the eigenvalue problem for matrices, which is what we need for any practical calculations.

The equation

$$\hat{O}|x_\alpha\rangle = \lambda_\alpha|x_\alpha\rangle, \qquad \alpha = 1,, 2, \ldots \tag{1}$$

is called the eigenvalue equation for the operator $\hat{O}$. Here λ_α, $\alpha = 1, 2, \ldots$, are unknown numbers and $|x_\alpha\rangle$, $\alpha = 1, 2, \ldots$, are unknown kets. The kets $|x_\alpha\rangle$ are called *eigenstates of* $\hat{O}$, or *eigenkets of* $\hat{O}$, or *eigenvectors of* $\hat{O}$. For each solution $|x_\alpha\rangle$, there is a number λ_α, called the *eigenvalue corresponding to the eigenket* $|x_\alpha\rangle$. The quantities λ_α and $|x_\alpha\rangle$ come always in pairs.

This kind of equation appears in many fields of physics. In quantum mechanics the operator $\hat{O}$ is likely to represent an observable O and this entails that $\hat{O}$ be a Hermitian operator. This guarantees that the spectrum (i.e. the set that includes all solutions λ_α, $\alpha = 1, 2, \ldots$) consists of real numbers. This must be the case because the theory was constructed so that the result of a measurement of O is one of the numbers in the spectrum. On physics grounds we expect therefore that there are as many eigenvalues as possible values for the observable. In the case of problems involving the states of the spin of one electron, there are only two eigenvalues. In general, eigenvalue problem for spins have a small number of eigenvalues. The spectrum of the harmonic oscillator or that of the particle in a box contains an infinite, but denumerable, number of values. Anything in between is possible if we are only interested in the discrete part of the spectrum (i.e. bound states).

Sometimes several eigenvectors correspond to the same eigenvalue. For example, we may have $\lambda_1 = \lambda_2 = \lambda_3$ while $|x_1\rangle \neq |x_2\rangle \neq |x_3\rangle$. When this happens we say that the eigenstates $|x_1\rangle$, $|x_2\rangle$, and $|x_3\rangle$ *are degenerate*.

The trivial solution $|0\rangle$ is physically meaningless and it is excluded.

The eigenvalue problem can be solved analytically only for a handful of systems. It is likely that all systems for which this is possible have been already examined and solved. For future work it is extremely probable that you will solve eigenvalue problems numerically. Except for problems involving collisions or photodissociation, we do this by converting the eigenvalue problem for a differential operator into an eigenvalue problem for the matrix representing the operator. The eigenvalue problem for the matrix is solved by using software provided by experts. Because of this, we focus here on understanding various aspects of the eigenvalue problem for matrices, and do not discuss numerical methods for solving it. When we need a numerical solution, we use `Mathematica`. More demanding calculations require the use of `Fortran`, `C`, or `C++` and of substantial computing power. These computer languages have libraries of programs designed to give you the eigenvalues and the eigenvectors when you give them the matrix.

Because most chemists are not familiar with the eigenvalue problem I examine here in great detail why an eigenvalue problem generates the results it does.

§ 4 *Some general observations.* Let us assume that we are interested in solving the eigenvalue equation

$$\hat{A}|\psi\rangle = \lambda|\psi\rangle \tag{2}$$

When the numerical indices in the eigenket and eigenvalues are not necessary I drop them (as I do in the equation given above). It is easy to see that if α is a complex number and $|\psi\rangle$ satisfies Eq. 2 then $|\eta\rangle \equiv \alpha|\psi\rangle$ satisfies the

same equation. For every given eigenvalue we have as many eigenstates as complex numbers α. The equation cannot tell us what value to use for α. We choose α so that $|\eta\rangle$ is normalized. We do this because pure states (the eigenstates of observables) are used to calculate probabilities, and in such a calculation all states involved must be normalized. Unless this is true the probability that something happens is not equal to one.

Next I remind you how we normalize a state. You need to know this because the library routines you may use may not be written to give normalized eigenvectors. Normalization means that we force the eigenstate $|\eta\rangle$ to satisfy

$$\langle \eta \,|\, \eta \rangle = 1 \tag{3}$$

Since $|\eta\rangle = \alpha|\psi\rangle$, Eq. 3 leads to

$$\alpha^* \alpha \langle \psi \,|\, \psi \rangle = 1 \tag{4}$$

or

$$\alpha^* \alpha = \frac{1}{\langle \psi \,|\, \psi \rangle} \tag{5}$$

Solving for α gives

$$\alpha = \frac{\mathrm{e}^{\mathrm{i}\phi}}{\sqrt{\langle \psi \,|\, \psi \rangle}} \tag{6}$$

where ϕ is a real number, which cannot be determined by solving Eq. 5. The ket

$$|\eta\rangle = \frac{\mathrm{e}^{\mathrm{i}\phi}}{\sqrt{\langle \psi \,|\, \psi \rangle}} |\psi\rangle \tag{7}$$

is normalized. The phase ϕ is still undetermined. However, we know that the factor $\exp[\mathrm{i}\phi]$ has no influence on any measurable quantity and therefore we can set it equal to 1.

Because $|\eta\rangle$ is equal to $|\psi\rangle$ multiplied with a number, if $|\psi\rangle$ is an eigenket corresponding to the eigenvalue λ then so is $|\eta\rangle$. If the system is in the state $|\eta\rangle$ and we measure A, which is the observable represented by $\hat{A}$, we are certain that the result is λ.

§ 5 *The orthogonality of the eigenstates.* If an operator represents an observable its eigenstates are pure states and we have shown in Chapter 2 that they must be orthonormal. What we found there is not erroneous but it needs to be refined. It is easy to show that the eigenvectors of a Hermitian operator that *correspond to different eigenvalues* must be orthogonal to each other. That is, if

$$\hat{A}|\eta\rangle = \lambda|\eta\rangle \quad \text{and} \quad \hat{A}|\eta'\rangle = \lambda'|\eta'\rangle \tag{8}$$

and

$$\lambda \neq \lambda' \tag{9}$$

then

$$\langle \eta \,|\, \eta' \rangle = 0 \tag{10}$$

To prove that this is true we note that for any Hermitian operator $\hat{A}$ we must have (see Chapter 4)

$$\langle \hat{A}\eta \,|\, \eta' \rangle = \langle \eta \,|\, \hat{A}\eta' \rangle \tag{11}$$

Using Eqs. 8 in Eq. 11 gives (because λ must be a real number)

$$(\lambda - \lambda')\langle \eta \,|\, \eta' \rangle = 0$$

When $\lambda \neq \lambda'$, we must have

$$\langle \eta \,|\, \eta' \rangle = 0,$$

which is what we wanted to prove.

What happens if some eigenstates have the same eigenvalues? Physicists call such eigenvalues 'degenerate', mathematicians are kinder and call them 'multiple'. If $|\eta\rangle$ and $|\eta'\rangle$ *belong to the same eigenvalue* (i.e. if $\hat{A}|\eta\rangle = \lambda|\eta\rangle$ and $\hat{A}|\eta'\rangle = \lambda|\eta'\rangle$) then the method used above to prove orthogonality fails because the proof assumed that $\lambda \neq \lambda'$, which is no longer true. When the eigenstates $|\eta\rangle$ and $|\eta'\rangle$ are degenerate they are no longer obligated to be orthogonal to each other. However, they must be orthogonal to the eigenvectors that belong to eigenvalues other than λ. Most numerical procedures for solving eigenvalue problems will provide non-orthogonal degenerate eigenvectors. We can collect all the kets that belong to the same eigenvalue λ and use the Gram-Schmidt procedure to convert this set of kets into a set of orthogonal kets. Then we can normalize these kets. The set of kets generated by this procedure belong to the eigenvalue λ, but now they are orthonormal. I will give an example of this procedure later in this chapter.

When we deal with degenerate eigenstates we have two sets of eigenvectors corresponding to the eigenvalue λ: the one produced when we solved the eigenvalue problem and the one obtained by orthnormalizing the first set. Which one is physically meaningful? The answer depends on the purpose for which we solved the eigenvalue problem. If we plan to use the eigenkets as a basis set then it is best to ensure that they are orthonormal. However if we want to analyze the physics of the system we need to do more work. We will analyze later a few examples of degeneracy and suggest that the eigenstates of an operator $\hat{A}$ are degenerate because they are simultaneously eigenstates of other operators which commute with $\hat{A}$. In a later chapter we

will prove that an operator has degenerate eigenvalues because it commutes with another operator that represents another observable.

§ 6 *The eigenvalue problem and the Fredholm alternative.* Write the eigenvalue problem for the operator $\hat{O}$ as

$$(\hat{O} - \lambda\hat{I})|\psi\rangle = |0\rangle \tag{12}$$

I have introduced here the ket $|0\rangle$ which is the equivalent of the number zero in the set of numbers. In other words, if $|\psi\rangle = |\phi\rangle$ then $|\psi\rangle = |\phi\rangle + |0\rangle$ is true for all kets $|\psi\rangle$ and $|\phi\rangle$; or (see Chapter 1) for every ket $|\psi\rangle$, $|\psi\rangle + |0\rangle = |\psi\rangle$. We need to use $|0\rangle$ because the difference $(\hat{O} - \lambda\hat{I})|\psi\rangle$ is a ket, not a number, and it is the ket obtained when we subtract two kets that are equal. Any operator acting on the zero ket gives $|0\rangle$ because otherwise acting on two equal kets with the same operator would not lead to equal results.

I am interested here in a Hermitian operator that represents an observable whose spectrum is discrete. Suppose we assume that the operator $(\hat{O} - \lambda\hat{I})^{-1}$ exists for all values of λ (i.e. for all real numbers). This leads to a small disaster because acting with $(\hat{O} - \lambda\hat{I})^{-1}$ on Eq. 12 gives

$$|\psi\rangle = (\hat{O} - \lambda\hat{I})^{-1}\,(\hat{O} - \lambda\hat{I})|\psi\rangle = (\hat{O} - \lambda\hat{I})^{-1}\,|0\rangle = |0\rangle \tag{13}$$

If $(\hat{O} - \lambda\hat{I})^{-1}$ exists for all values of λ then the operator $\hat{O}$ has only one eigenstate, namely $|0\rangle$, which is not a physically acceptable state. However the operator that represents an observable must have eigenstates that are not $|0\rangle$. We conclude that there must be values of λ for which $(\hat{O} - \lambda\hat{I})^{-1}$ does not exist and these must be the eigenvalues of $\hat{O}$.

It is no hard to prove this statement. Remember that we can write

$$\hat{O} = \sum_{\alpha} |a_\alpha\rangle \, \lambda_\alpha \, \langle a_\alpha| \tag{14}$$

where $|a(\alpha)\rangle$ are the eigenvectors and $\lambda(\alpha)$ the corresponding eigenvalues of $\hat{O}$. Since $(\hat{O} - \lambda\hat{I})^{-1}$ is a function of $\hat{O}$, we can write it as

$$(\hat{O} - \lambda\hat{I})^{-1} = \sum_{\alpha} |a_\alpha\rangle \, \frac{1}{\lambda_\alpha - \lambda} \, \langle a_\alpha| \tag{15}$$

Because $\hat{O}$ is Hermitian, all the eigenvalues λ_α are real numbers. Whenever $\lambda = \lambda_\alpha$, the term $1/(\lambda_\alpha - \lambda)$ becomes infinite and $(\hat{O} - \lambda\hat{I})^{-1}$ does not exist. This proves the assertion that if λ is an eigenvalue of $\hat{O}$ then $(\hat{O} - \lambda\hat{I})^{-1}$ does not exist and that all eigenvalues of $\hat{O}$ are accounted for by this statement.

Consider now the equation

$$(\hat{O} - \lambda\hat{I})|\psi\rangle = |\phi\rangle \tag{16}$$

where $|\phi\rangle$ is known but $|\psi\rangle$ is unknown. If λ is such that $(\hat{O} - \lambda\hat{I})^{-1}$ exists then Eq. 16 has the solution

$$|\,\psi\rangle = (\hat{O} - \lambda\hat{I})^{-1}|\phi\rangle \tag{17}$$

This is true for all values of λ that *are not eigenvalues of* $\hat{O}$.

The set of real numbers is thus divided into two non-overlapping sets.

1. The eigenvalues λ_α, $\alpha = 1, 2, \ldots$ of $\hat{O}$, for which
   $(\hat{O} - \lambda\hat{I})|\psi\rangle = |0\rangle$ has solutions different from $|0\rangle$
   and
   $(\hat{O} - \lambda\hat{I})|\psi\rangle = |\phi\rangle$ does not have a solution for nonzero $|\phi\rangle$.

2. All the other real numbers λ, for which
   $(\hat{O} - \lambda\hat{I})|\psi\rangle = |0\rangle$ has only the solution $|0\rangle$
   and
   $(\hat{O} - \lambda\hat{I})|\psi\rangle = |\phi\rangle$ has nonzero solutions for $|\,\psi\rangle$.

This division into classes is sometimes call the Fredholm alternative.[1] It creates an interesting situation: we can find the eigenvalues λ_α of $\hat{O}$ by requiring that the real number λ has the property that

$$(\hat{O} - \lambda\hat{I})|\psi\rangle = |\phi\rangle \tag{18}$$

does not have a solution. We will use this in what follows.

§ 7 *The eigenvalue problem in matrix representation: a review.* In Chapter 5, I have shown that by choosing an N-dimensional orthonormal basis set, we can represent any ket by an N-dimensional vector and any operator by an $N \times N$ matrix. These connections are isomorphisms: all relationships between kets are translated into identical relationships for N-dimensional vectors and all operator relationships are maintained as matrix relationships. Here I review these connections, since calculations with matrices and vectors are more efficient and simpler than using differential operators.

Let $|a_1\rangle$, $|a_2\rangle$, ..., $|a_n\rangle$, ... be a complete orthonormal basis set. This set need not be a set of pure states of an observable; we could use any complete basis set that a mathematician can construct. Orthonormality means that

$$\langle a_i \,|\, a_j\rangle = \delta_{ij} \text{ for all } i, j \tag{19}$$

[1] Fredholm developed this theory for linear integral operators but the name is now used for any linear operator.

Because the set is complete, we have

$$\hat{I} \simeq \sum_{n=1}^{N} |a_n\rangle \langle a_n| \tag{20}$$

Eq. 20 ignores the continuous spectrum and truncates the infinite sum in the completeness relation.

We can act with each $\langle a_m|$, $m = 1, 2, \ldots, N$, on the eigenvalue equation $\hat{O}|\psi\rangle = \lambda|\psi\rangle$ to turn it into the following N equations:

$$\langle a_m \,|\, \hat{O} \,|\, \psi\rangle = \lambda \langle a_m \,|\, \psi\rangle, \quad m = 1, 2, \ldots, N \tag{21}$$

Now insert $\hat{I}$, as given by Eq. 20, between $\hat{O}$ and $|\psi\rangle$ to obtain

$$\sum_{n=1}^{N} \langle a_m \,|\, \hat{O} \,|\, a_n\rangle \langle a_n \,|\, \psi\rangle = \lambda \langle a_m \,|\, \psi\rangle, \quad m = 1, 2, \ldots, N \tag{22}$$

The complex numbers $\langle a_n \,|\, \psi\rangle$ are the coordinates of $|\psi\rangle$ in the $\{|a_n\rangle\}_{n=1}^{N}$ representation, and I denote them by ψ_n, which is the usual notation in linear algebra. If we know them, we can write $|\psi\rangle$ as

$$|\psi\rangle = \sum_{n=1}^{N} |a_n\rangle \, \psi_n \tag{23}$$

Following again the customs of linear algebra, we use the notation

$$O_{mn} \equiv \langle a_m \,|\, \hat{O} \,|\, a_n\rangle \tag{24}$$

With this notation, the eigenvalue equation, Eq. 22, becomes

$$\sum_{n=1}^{N} O_{mn} \, \psi_n = \lambda \psi_m, \quad m = 1, 2, \ldots, N \tag{25}$$

The sum in Eq. 25 is the rule by which the matrix O, having the elements O_{mn}, acts on the vector ψ, having the coordinates ψ_n. This equation is the *eigenvalue problem for the matrix* O and it is often written as

$$O\psi = \lambda\psi \tag{26}$$

(which resembles the operator equation $\hat{A}|\psi\rangle = \lambda|\psi\rangle$) or as

$$\begin{pmatrix} O_{11} & O_{12} & \cdots & O_{1N} \\ O_{21} & O_{22} & \cdots & O_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ O_{N1} & O_{N2} & \cdots & O_{NN} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} = \lambda \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} \tag{27}$$

Eqs. 26 and 27 are different ways of writing Eq. 25, which, in turn, is the representation of the equation $\hat{O}|\psi\rangle = \lambda|\psi\rangle$ in the basis set $\{|a_1\rangle, \ldots, |a_N\rangle\}$.

Normally in quantum mechanics we *choose* the basis set $\{|a_1\rangle, \ldots, |a_n\rangle\}$ and *know* the operator $\hat{A}$. This means that we can calculate $A_{mn} = \langle a_m \,|\, \hat{A} \,|\, a_n\rangle$. We do not know $|\psi\rangle$ or λ, and our mission is to find them from $\hat{A}|\psi\rangle = \lambda|\psi\rangle$. We have converted this operator equation, by using the orthonormal basis set $\{|a_1\rangle, \ldots, |a_N\rangle\}$, into the matrix equation Eq. 25 (or Eq. 26 or Eq. 27). In Eq. 25, we know A_{mn} but we do not know $\psi = \{\psi_1, \psi_2, \ldots, \psi_N\}$ or λ. We must calculate them from Eq. 25. Once we know $\psi_1, \psi_2, \ldots, \psi_N$, we can calculate $|\psi\rangle$ from Eq. 23 .

Here are the steps of the procedure:

1. Pick a basis set $|\, a_\alpha\rangle$ with $\alpha = 1, 2, 3, \ldots, N$

2. Calculate the matrix elements $\langle a_\alpha \,|\, \hat{O} \,|\, a_\gamma\rangle$ for $\alpha = 1, 2, 3, \ldots, N$ and $\gamma = 1, 2, 3, \ldots, N$

3. Use a computer to find the eigenvectors and the eigenvalues of the matrix.

4. You can use the eigenvectors of the matrix in Eq. 25 to calculate the eigenkets and use them to calculate any observable.

Without going into details (which are examined later), I warn you that the accuracy depends on your ability to design a good basis set and pick the right value for N, and that for fixed basis set and N, you are likely to get the small eigenvalue (the ground state) more accurately than the larger values (the excited states). Moreover, this method cannot be used for solving the eigenvalue problem for processes involving collisions or photodissociation.

The methods used in what follows are not limited to quantum mechanics: they can use to solve any eigenvalue problem for the linear operators in any branch of physics (e.g. acoustics, elasticity, or electrodynamics)

§ 8 *The Fredholm alternative for linear algebra.* As I have pointed out, the conversion of kets to vectors and of operators to matrices are isomorphisms. This means that these conversions preserve relationships: an equation for an operator and a ket becomes the same equation for the corresponding matrix and vector. In particular, the eigenvalue equation for the matrix O is

$$(O - \lambda I)\psi = \mathbf{0} \tag{28}$$

where $\mathbf{0}$ is an N-dimensional vector whose components are all equal to zero. This equation has a solution only for those values of λ for which

$$(O - \lambda I)\psi = \phi \tag{29}$$

does not have a solution. Both equations are used in quantum mechanics but we are focusing now on the eigenvalue equation Eq. 28. This has a nonzero solution only if $O - \lambda I$ does not have an inverse.

I will prove in what follows that the inverse does not exist if the determinant of $O - \lambda I$ is zero. To proceed I define first the determinant of a matrix

to be

$$\det[A] = a_1 a_2 a_3 \cdots a_N \tag{30}$$

where a_i, $i = 1, 2, \ldots, N$ are the eigenvalues of A. This is different from the definition of the determinant given in most books but it is as good as any other definition. In the Appendix of Chapter 4 I listed many properties of the determinant, but this definition is the only thing we need right now. From operator theory we know that the inverse of $O - \lambda I$ is

$$(O - \lambda I)^{-1} = \sum_{i=1}^{N} |o_i\rangle \frac{1}{o_i - \lambda} \langle o_i| \tag{31}$$

where o_i are the eigenvalues of O and $|o_i\rangle$ are a basis set of eigenvectors. The determinant of $O - \lambda I$ is

$$\det[O - \lambda I] = (o_1 - \lambda)(o_2 - \lambda)...(o_N - \lambda) \tag{32}$$

If λ is equal to one of the eigenvalues o_i, the inverse og $(O - I\lambda)$ does not exist and $\det[O - \lambda I] = 0$. The reverse is also true: if the determinant is zero the inverse does not exist. Another way of stating this is that the eigenvalues of O are the roots of the equation

$$\det[O - I\lambda] = 0 \tag{33}$$

From Eq. 32 we see that $\det[O - I\lambda]$ is an N-th order polynomial in λ and that the eigenvalues of the matrix O are the roots of this polynomial. $\det[O - I\lambda]$ is called *the characteristic polynomial of matrix O*. Because the matrix is Hermitian its eigenvalues must be real numbers, which means that all roots of the characteristic polynomial are real numbers[2].

[2]If you are a gambling person bet someone that you can generate at will polynomials of arbitrary (but finite) order that have only real roots. What do you do to win the bet?

§ 9 *Summary.* If $\hat{A}$ is the operator describing an observable we can find its spectrum and its pure states by picking an orthonormal basis set, calculating the matrix elements of $\hat{A}$, and solving the eigenvalue problem for the matrix A. The "quantization" of the values that the observable can take occurs because $(A - \lambda I)\psi = \mathbf{0}$ has solutions only for those special values of λ for which the matrix $A - \lambda I$ does not have an inverse. These are the values of λ for which $\det[A - \lambda I] = 0$. Note that the spectrum of A depends only on the matrix A, as it should.

If we use N functions in our orthonormal, complete basis set, the matrix A can have only N eigenvalues. They are not necessarily distinct, because some roots of the characteristic polynomial might be equal to each other.

In abstract this procedure seems straightforward. However, choosing a basis set that is appropriate for a given physical system is an art that improves with experience. We will discuss various issues related to the choice of basis set in future chapters when we apply this method.

Section 9.2. A Theorem Regarding the Eigenvalue Problem

§ 10 *Statement.* I derive here a theorem that is often used when establishing various relationships or when performing computations. The theorem involves two operators, $\hat{H}$ and $\hat{U}$, and it is most useful when $\hat{H}$ is Hermitian and $\hat{U}$ is unitary. However, I make those assumptions below only when necessary.

Before formulating and proving the theorem, I will revisit the notation. Most of the time, I have used the notation $\hat{H}|x\rangle = \lambda|x\rangle$ or $\hat{H}|x_i\rangle = \lambda|x_i\rangle$ for

the eigenvalue problem for operator $\hat{H}$. I am changing that now to

$$\hat{H}|x(i)\rangle = \lambda(i)|x(i)\rangle \tag{34}$$

I wish to use $|x(i)\rangle$ rather than $|x_i\rangle$ because I will work in the ket space and in $\boldsymbol{\ell}^2$. In $\boldsymbol{\ell}^2$ the ket $|x(i)\rangle$ is represented by the vector

$$x(i) = \{x_1(i), x_2(i), x_3(i), \ldots, x_N(i)\} = \{x_\alpha(i)\}_{\alpha=1}^N \tag{35}$$

Here

$$x_\alpha(i) = \langle b_\alpha \,|\, x(i)\rangle \tag{36}$$

where $|b_\alpha\rangle$, $\alpha = 1, \ldots, N$ is some orthonormal basis set used to define $\boldsymbol{\ell}^2$. I will use lower-case roman letters to label kets (as in $|x(i)\rangle$) and lower-case greek letters to label the components of the vector representing a ket in $\boldsymbol{\ell}^2$. Greek letters will also label the matrix elements. For example, the eigenvalue equation Eq. 34 will be written in $\boldsymbol{\ell}^2$ as

$$H_{\alpha\beta}x_\beta(i) = \lambda(i)x_\alpha(i) \tag{37}$$

In this notation, $x_2(6)$ is the second component of the sixth eigenvector of H.

We can now proceed to state the theorem.

1. If $\hat{U}$ is an operator that has an inverse then $\hat{U}^{-1}\hat{H}\hat{U}$ has the same eigenvalues as $\hat{H}$:

   $$\hat{U}^{-1}\hat{H}\hat{U}|e(i)\rangle = \lambda(i)|e(i)\rangle \tag{38}$$

   if

   $$\hat{H}|x(i)\rangle = \lambda(i)|x(i)\rangle \tag{39}$$

Note that the eigenvalues in these two equations are the same. In addition, the eigenvectors $|e(i)\rangle$ are related to $|x(i)\rangle$ through

$$|x(i)\rangle = \hat{U}|e(i)\rangle \tag{40}$$

2. If the operator $\hat{U}$ in item 1 is unitary and the operator $\hat{H}$ is Hermitian, then $\hat{U}^{-1}\hat{H}\hat{U}$ is Hermitian, and both $\{|x(i)\rangle\}_{i=1}^{N}$ and $\{|e(i)\rangle\}_{i=1}^{N}$ are complete, orthonormal basis sets.

3. Continuing the notation in item 1, if $\{|e_\alpha(i)\rangle\}_{\alpha=1}^{N}$ is used as a basis set then the matrices U, U^{-1}, and H corresponding to the operators $\hat{U}$, $\hat{U}^{-1}$, and $\hat{H}$ satisfy

$$U^{-1}HU = \Lambda \tag{41}$$

where Λ *is a diagonal matrix* whose diagonal elements are the eigenvalues of H. In symbols,

$$\Lambda_{ij} = \lambda(i)\delta_{ij} \tag{42}$$

with $\lambda(i)$, $i = 1, 2, \ldots, N$ the eigenvalues of H.

4. If the eigenvectors $x(i), i = 1, 2, \ldots, N$ of H are known and their components are denoted by $x_\alpha(i)$ (that is, $x(i) = \{x_1(i), x_2(i), \ldots, x_N(i)\}$) then the matrix U is

$$U = \begin{pmatrix} x_1(1) & x_1(2) & \cdots & x_1(N) \\ x_2(1) & x_2(2) & \cdots & x_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_N(1) & x_N(2) & \cdots & x_N(N) \end{pmatrix} \tag{43}$$

U is constructed so that each column is a normalized eigenvector of H. If some eigenvectors are degenerate they should be orthonormalized (by the Gram-Schmidt procedure) before being used in Eq. 43.

§ 11 *Proof.* If $|x(i)\rangle$ is any eigenvector corresponding to some eigenvalue $\lambda(i)$ of $\hat{H}$, so that

$$\hat{H}|x(i)\rangle = \lambda(i)|x(i)\rangle \tag{44}$$

then define $|e(i)\rangle$ through

$$|x(i)\rangle \equiv \hat{U}|e(i)\rangle \tag{45}$$

At this point $\hat{U}$ can be any operator that has an inverse. Introduce $|x(i)\rangle$ given by Eq. 45 into Eq. 44, act with $\hat{U}^{-1}$ on the resulting equation, and use $\hat{U}^{-1}\hat{U} = \hat{I}$. You get

$$\hat{U}^{-1}\hat{H}\hat{U}|e(i)\rangle = \lambda(i)|e(i)\rangle \tag{46}$$

This means that $\hat{U}^{-1}\hat{H}\hat{U}$ has the same eigenvalues as $\hat{H}$

More is gained if $\hat{U}$ is unitary. In that case, we know that $\hat{U}^{-1}$ exists and in fact $\hat{U}^{-1} = \hat{U}^{\dagger}$,and we have

$$\hat{U}^{-1}\hat{H}\hat{U} = \hat{U}^{\dagger}\hat{H}\hat{U}$$

If $\hat{H}$ is Hermitian, then $\hat{U}^{-1}\hat{H}\hat{U}$ is also Hermitian (use $(\hat{A}\hat{B})^{\dagger} = \hat{B}^{\dagger}\hat{A}^{\dagger}$ and $\hat{H}^{\dagger} = \hat{H}$). In addition, if $\hat{U}$ is unitary and $\{|x(i)\rangle\}_{i=1}^{N}$ is a complete, orthonormal basis set, then $\{|e(i)\rangle\}_{i=1}^{N}$ must also be a complete, orthonormal basis set.

Note that if $\hat{H}$ is Hermitian but has degenerate eigenstates, they are not automatically orthonormal but can be always converted to an orthonormal set by using the Gram-Schmidt method. This proves item 2.

We can now move on to prove item 3. Start with Eq. 46 and act on it with $\langle e(j)|$, then use orthonormality:

$$\langle e(j) \,|\, \hat{U}^{-1}\hat{H}\hat{U} \,|\, e(i)\rangle = \lambda(i)\delta_{ji} \tag{47}$$

By inserting the completeness relation $\sum_k |e(k)\rangle \langle e(k)| = \hat{I}$ between $\hat{U}^{-1}$ and $\hat{H}$ and the completeness relation $\sum_p |e(p)\rangle \langle e(p)| = \hat{I}$ between $\hat{H}$ and $\hat{U}$, and using $\langle e(n) \,|\, \hat{O} \,|\, e(m)\rangle = O_{nm}$, we obtain

$$U^{-1}_{jk} H_{kp} U_{pi} = \Lambda_{ji} \tag{48}$$

This is Eq. 41, which we wanted to prove.

Let us prove item 4 by finding the matrix U. Because we defined

$$\hat{U}|e(i)\rangle = |x(i)\rangle \tag{49}$$

we have (multiply Eq. 49 with $\langle e(j)|$)

$$U_{ji} \equiv \langle e(j) \,|\, \hat{U} \,|\, e(i)\rangle = \langle e(j) \,|\, x(i)\rangle \tag{50}$$

But $\langle e(j) \,|\, x(i)\rangle \equiv x_j(i)$ is the j-th "coordinate" of the eigenvector $x(i)$. Therefore Eq. 50 is

$$U_{ji} = x_j(i)$$

Since the index i in U_{ji} labels the columns of the matrix U, we have

$$U = \begin{pmatrix} x_1(1) & x_1(2) & \cdots & x_1(N) \\ x_2(1) & x_2(2) & \cdots & x_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_N(1) & x_N(2) & \cdots & x_N(N) \end{pmatrix}$$

The columns of U are the eigenvectors of $\hat{H}$, as stated in the theorem.

It is important to note the following. The matrix U is not unique. We defined it by using the orthonormal basis set $|e(i)\rangle$. But there is an infinite number of orthonormal basis sets. Acting with any unitary operator $\hat{Q}$ on $|e(i)\rangle$ generates another basis set $|y(i)\rangle \equiv \hat{Q}|e(i)\rangle$. We could have used $|y(i)\rangle$ to define $\hat{U}$ (i.e. $|x(i)\rangle \equiv \hat{U}|y(i)\rangle$) and the proof will work as before. At this point you might protest: U must be formed by using the eigenvectors of H as columns. Doesn't that pin down U? No. The eigenvectors of H depend on the basis set also. I will build U in the basis set $\{|y(i)\rangle\}_{i=1}^{N}$ by using the eigenvectors of H in *the same basis set.* Everything seems to change with the basis set but the theorem stays the same.

The `Mathematica` file `linear algebra for quantum mechanics` shows you how to use `Mathematica` to solve problems in linear algebra. In particular, it tests the theorem just derived.

Exercise 1 1. Use the file `Linear algebra for quantum mechanics` to generate a 4×4 Hermitian matrix H and two 4×4 unitary matrices U and Q.

2. Calculate the eigenvectors and eigenvalues of H. Call the eigenvalues λ[i], $i = 1, 2, 3, 4$, and the eigenvectors h[i], $i = 1, 2, 3, 4$ (look at the function `Eigensystems` in the `Mathematica` file).

3. Calculate the eigenvectors and eigenvalues of $U^{-1}HU$ and call them uhu[i] and uλu[i]. Test that the eigenvalues of $U^{-1}HU$ are the same as those of H.

4. Calculate the eigenvectors and eigenvalues of $Q^{-1}HQ$ and call them

qhq[i] and qλq[i]. Test that the eigenvalues of $Q^{-1}HQ$ are the same as those of H.

5. Test that the eigenvectors h[i] of H are the same (up to a phase factor) as Q.qhq[i]. Remember that they are the same if they differ by a phase factor.

6. Test that the matrix W formed by placing the eigenvectors of H in columns is unitary and that $W^{-1}HW$ is a diagonal matrix for which the diagonal elements are the eigenvalues of H.

Section 9.3. Nondegenerate Eigenvalues

§ 12 *Introduction.* The previous two section describe what you need to know, in general, about the eigenvalues of a matrix. In what follows I solve the eigenvalue problem for a particular 3×3 matrix. The purpose is to show in detail, by example, the connection between the eigenvalue problem and homogeneous linear equations as well as the role of normalization for obtaining unique eigenvectors. I will also examine an example of eigenvalue problem that has degenerate eigenvalues. If you are comfortable with the general theory you could skip this section, but I wouldn't.

§ 13 *A three-dimensional example.* Let us examine the eigenvalue problem for the matrix

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 1.2 & 3.1 \\ 4 & 3.1 & 4 \end{pmatrix} \tag{51}$$

The characteristic equation is (see Eq. 33)

$$\det \begin{pmatrix} 3-\lambda & 2 & 4 \\ 2 & 1.2-\lambda & 3.1 \\ 4 & 3.1 & 4-\lambda \end{pmatrix} = 0 \tag{52}$$

If you find tedious algebra soothing, you can calculate this determinant by hand. I prefer to use `Mathematica`, and the result is (see Section 2, Cell 1 in the file `WorkBook9 The eigenvalue problem.nb`)

$$-\lambda^3 + 8.2\lambda^2 + 9.21\lambda - 0.03 = 0 \tag{53}$$

The characteristic polynomial of the matrix A is a third-order polynomial because we used a three-dimensional basis set and therefore A is a 3×3 matrix.

The roots of the characteristic equation are the eigenvalues of A. They are (see Section 2, Cell 1 of `WorkBook9`)

$$\left.\begin{aligned} \lambda_1 &= -1.00391 \\ \lambda_2 &= 0.0032479 \\ \lambda_3 &= 9.20066 \end{aligned}\right\} \tag{54}$$

The eigenvalues must be real numbers because A is Hermitian.

Exercise 2 Use `Mathematica` to check that $\det A = \lambda_1\lambda_2\lambda_3$.

How do we find the eigenvectors? For the eigenvector $\psi(1)$ corresponding to the eigenvalue λ_1, insert λ_1 in the equation $A\psi = \lambda\psi$ to obtain

$$A\psi(1) - \lambda_1\psi(1) = 0, \tag{55}$$

The list $\psi(1) = \{\psi_1(1), \psi_2(1), \psi_3(1)\}$ is the eigenvector corresponding to the eigenvalue λ_1. If I insert λ_2 in $A\psi = \lambda\psi$, I get a different equation

$$A\psi(2) - \lambda_2\psi(2) = 0 \tag{56}$$

for a different eigenvector, $\psi(2)$, corresponding to the eigenvalue λ_2. Similarly,

$$A\psi(3) - \lambda_3\psi(3) = 0 \tag{57}$$

gives the third eigenvector $\psi(3)$. Because I have three values for λ, I have three distinct equations for ψ, and their solutions are $\psi(1)$, $\psi(2)$, and $\psi(3)$.

To find the eigenvector $\psi(1)$, we must solve the homogeneous system of equations

$$(A - \lambda_1 I)\psi(1) = \begin{pmatrix} 3-\lambda_1 & 2 & 4 \\ 2 & 1.2-\lambda_1 & 3.1 \\ 4 & 3.1 & 4-\lambda_1 \end{pmatrix} \begin{pmatrix} \psi_1(1) \\ \psi_2(1) \\ \psi_3(1) \end{pmatrix} = \mathbf{0}$$

Performing the matrix-vector multiplication, we find the three equations

$$\begin{aligned} (3-\lambda_1)\psi_1(1) + 2\psi_2(1) + 4\psi_3(1) &= 0 && (58) \\ 2\psi_1(1) + (1.2-\lambda_1)\psi_2(1) + 3.1\psi_3(1) &= 0 && (59) \\ 4\psi_1(1) + 3.1\psi_2(1) + (4-\lambda_1)\psi_3(1) &= 0 && (60) \end{aligned}$$

We seem to have a difficulty here. If $\{\psi_1(1), \psi_2(1), \psi_3(1)\}$ is a solution then $\{\beta\psi_1(1), \beta\psi_2(1), \beta\psi_3(1)\}$ is also a solution. This is not a real problem, however, since, as noted earlier, we can fix β by requiring that the eigenvector be normalized. The normalized vector is the only meaningful eigenvector in quantum mechanics.

I can look at Eqs. 58–60 and see right away that $\psi = \{0, 0, 0\}$ is a solution. This is not physically acceptable, because it cannot represent a state of the system. Are there other solutions besides $\{0, 0, 0\}$? The answer is yes. Here is how we find them.

Divide Eqs. 58–60 by $\psi_3(1)$, we obtain

$$(3 - \lambda_1)\frac{\psi_1(1)}{\psi_3(1)} + 2\frac{\psi_2(1)}{\psi_3(1)} = -4 \tag{61}$$

$$2\frac{\psi_1(1)}{\psi_3(1)} + (1.2 - \lambda_1)\frac{\psi_2(1)}{\psi_3(1)} = -3.1 \tag{62}$$

$$4\frac{\psi_1(1)}{\psi_3(1)} + 3.1\frac{\psi_2(1)}{\psi_3(1)} = \lambda_1 - 4 \tag{63}$$

The homogeneous system of three equations and three unknowns has become an *inhomogeneous system of three equations and two unknowns* $\psi_1(1)/\psi_3(1)$ *and* $\psi_2(1)/\psi_3(1)$. We seem to have two new difficulties here: we have one equation too many and we can only calculate the ratios $\psi_1(1)/\psi_3(1)$ and $\psi_2(1)/\psi_3(1)$. We will fix $\psi_3(1)$ by imposing the condition that the eigenvectors must be orthonormal (because they are not degenerate). Having too many equations cannot be helped. However, I can pick any two equations out of the three and solve them for $\psi_1(1)/\psi_3(1)$ and $\psi_2(1)/\psi_3(1)$. The theory is consistent only I get the same eigenvector regardless of which pair of equations I pick. We will see shortly that this is the case. I will also show that this consistency is achieved only when λ_1 is an eigenvalue.

I pick the first two equations (Eq 61 and 62) and solve them in `WorkBook9`, Section 2, Cell 2. The result is

$$\psi_1(1) = -0.542\,\psi_3(1) \tag{64}$$

$$\psi_2(1) = -0.914\,\psi_3(1) \tag{65}$$

so the eigenvector $\psi(1)$ is

$$\begin{aligned}\psi(1) &= \{-0.542\,\psi_3(1), -0.914\,\psi_3(1), \psi_3(1)\} \\ &= \psi_3(1)\{-0.542, -0.914, 1\} \end{aligned} \tag{66}$$

Of course, $\psi_3(1)$ is not yet determined. The unknown $\psi_3(1)$ is a multiplicative constant, which I determine by imposing the normalization requirement

$$1 = \langle\psi(1)\,|\,\psi(1)\rangle = 2.13\,|\psi_3(1)|^2 \tag{67}$$

Eq. 67 has two solutions, 0.685 and -0.685. They are equivalent, since they differ by a factor whose absolute value is 1; such "phase factors" make no difference in any calculation of an observable quantity. I use the positive root:

$$\psi_3(1) = 0.685 \tag{68}$$

and with this, I have

$$\psi(1) = 0.685\{-0.542, -0.914, 1\} = \{-0.371, -0.626, +0.685\} \tag{69}$$

As I already said, the theory is consistent only if any other pair of two equations gives the same eigenvectors. In `WorkBook9` I showed that this is true. This must be so but it seems rather miraculous that it happens. However, no miracle is involved. Let us assume that we did not calculate the eigenvalue λ and left it undetermined in Eqs. 58–60. We can take Eqs. 58 and 59 and solve for the ratios $\psi_1(1)/\psi_3(1)$ and $\psi_2(1)/\psi_3(1)$. We obtain two solutions that depend on λ. Next we take Eqs. 58 and 60 and solve for $\psi_1(1)/\psi_3(1)$ and $\psi_2(1)/\psi_3(1)$, obtaining two new solutions for those ratios. But the ratios obtained in the first calculation *must* be equal to the ratios

obtained from the second calculation. This equality gives me an equation for λ. The values of λ obtained in this way are the eigenvalues.

Exercise 3 Implement the recipe outlined above for the homogeneous system

$$(3.2 - \lambda)\alpha + 2.4\beta = 0 \quad (70)$$

$$2.4\alpha + (3.6 - \lambda)\beta = 0 \quad (71)$$

Solve each equation for α/β and demand that the ratio (α/β) obtained from Eq. 70 be equal to the ratio obtained from Eq. 71. This gives you an equation for λ; call it EQ. Then calculate the characteristic polynomial of the matrix and test whether it is the same as EQ. After that, go ahead and find the eigenvalues and the eigenvectors.

Section 9.4. Degenerate Eigenvalues

§ 14 *Introduction.* The eigenvalues of a matrix A are the roots of the characteristic polynomial $\det[A - \lambda I]$. There is no law that prevents some of these roots from being equal. When two or more eigenvalues are equal, we say that they are *degenerate* and the corresponding eigenvectors are also called *degenerate*. This name suggests a certain moral depravity for these eigenvalues and I am sure that they find it offensive. I don't know who introduced this terminology. Mathematicians are more polite and use instead the terms 'multiple' and 'single' eigenvalues, which is simpler and friendlier.

One would think that finding operators that have identical eigenvalues should be as rare as finding by accident polynomials with multiple roots. But this is not so. You will find out in future chapters that physics requires in some cases that the eigenvalues be degenerate. You can see some examples below.

§ 15 *Particle in a cubic box with infinite potential-energy walls.* One of the simplest models in quantum mechanics is a particle in a rectangular parallelepiped whose walls are infinitely repulsive.

The total energy of the particle is discrete and is given by (see H. Metiu, *Physical Chemistry, Quantum Mechanics*, page 104)

$$E_{n,j,k} = \frac{\hbar^2\pi^2}{2m}\left(\frac{n}{L_x}\right)^2 + \frac{\hbar^2\pi^2}{2m}\left(\frac{j}{L_y}\right)^2 + \frac{\hbar^2\pi^2}{2m}\left(\frac{k}{L_z}\right)^2 \tag{72}$$

where n, j, k can take any of the values

$$n, j, k = 1, 2, 3, \ldots \tag{73}$$

Here L_x, L_y, and L_z are the lengths of the box edges.

The ket representing the pure states (the eigenstates of $\hat{H}$) is denoted by $|n, j, k\rangle$. For example, the symbol $|2, 1, 1\rangle$ means the state with $n = 2$, $j = 1$, $k = 1$, whose energy is

$$E_{2,1,1} = \frac{\hbar^2\pi^2}{2m}\left[\left(\frac{2}{L_x}\right)^2 + \left(\frac{1}{L_y}\right)^2 + \left(\frac{1}{L_z}\right)^2\right] \tag{74}$$

Exercise 4 Write down the energy of the states $|1, 2, 1\rangle$ and $|1, 1, 2\rangle$.

If $L_x \neq L_y \neq L_z$, the states are not degenerate (except if accidentally $\left(\frac{n}{L_x}\right)^2 + \left(\frac{k}{L_y}\right)^2 + \left(\frac{m}{L_z}\right)^2 = \left(\frac{n'}{L_x}\right)^2 + \left(\frac{k'}{L_y}\right)^2 + \left(\frac{m'}{L_z}\right)^2$ in which case $|n, j, m\rangle$ is degenerate with $|n', j', m'\rangle$).

But consider the case of a cubic box for which

$$L_x = L_y = L_z = L \tag{75}$$

In this case the states $|2, 1, 1\rangle$, $|1, 2, 1\rangle$, and $|1, 1, 2\rangle$ have the same energy, equal to

$$\frac{\hbar^2\pi^2}{2mL^2}\left(2^2 + 1 + 1\right) \tag{76}$$

They are *degenerate* states.

Note the connection between degeneracy and symmetry. If $L_x \neq L_y \neq L_z$, there is no degeneracy, except perhaps by accident. However, if $L_x = L_y = L_z$, the box is symmetric and almost all states are degenerate.

Exercise 5 Is the state $|1, 1, 1\rangle$ degenerate? How about $|3, 2, 1\rangle$? Enumerate the states having the same energy as $|3, 2, 1\rangle$.

In the Schrödinger representation, the state $|n, j, k\rangle$ is (see Metiu, Eq. 8.31)

$$\begin{aligned} \langle x, y, z \,|\, n, j, k\rangle &\equiv \psi_{n,j,k}(x, y, z) \\ &= \sqrt{\frac{2}{L_x}}\sin\left(\frac{n\pi x}{L_x}\right)\sqrt{\frac{2}{L_y}}\sin\left(\frac{j\pi y}{L_y}\right)\sqrt{\frac{2}{L_z}}\sin\left(\frac{k\pi z}{L_z}\right) \end{aligned} \tag{77}$$

For the example of the cube,

$$\psi_{2,1,1}(x, y, z,) = \sqrt{\frac{8}{L^3}}\,\sin\left(\frac{2\pi x}{L}\right)\,\sin\left(\frac{\pi y}{L}\right)\,\sin\left(\frac{\pi z}{L}\right) \tag{78}$$

and

$$\psi_{1,2,1}(x,y,z) = \sqrt{\frac{8}{L^3}} \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{2\pi y}{L}\right) \sin\left(\frac{\pi z}{L}\right) \tag{79}$$

A similar equation holds for $\psi_{1,1,2}(x,y,z)$. These three states have the same energy *but they are different states*. What is the physical difference between them?

Let us look at the kinetic energy in the x-direction:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi_{2,1,1}(x,y,z) = \frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2 \psi_{2,1,1}(x,y,z) \tag{80}$$

I obtained the second line from the first by taking the second derivative of $\psi_{2,1,1}(x,y,z)$ given by Eq. 78. We see that $|2,1,1\rangle$ is an eigenstate of $\hat{K}_x$, with the eigenvalue $\frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2$. In the same way, you can show that $|2,1,1\rangle$ is an eigenstate of $\hat{K}_y$ and $\hat{K}_z$, with the eigenvalues $\frac{\hbar^2}{2m}\left(\frac{1\pi}{L}\right)^2$. The total energy is the sum of those three kinetic energies. In the state $|2,1,1\rangle$, the particle has higher kinetic energy in the x-direction. We can analyze in the same way the states $|1,2,1\rangle$ and $|1,1,2\rangle$. Table 1 gives the result. The three degenerate states have the high (i.e. $\frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2$) kinetic energy in *different directions*.

Can we distinguish these three states by some measurement? We can. The particle in the state $|2,1,1\rangle$ will emit a photon in a different direction than will a particle in $|1,2,1\rangle$ or $|1,1,2\rangle$.

The degenerate states $|2,1,1\rangle$, $|1,2,1\rangle$, and $|1,1,2\rangle$ are different physical states, even though they have the same energy.

We have also found a curious thing. The degenerate states $|2,1,1\rangle$, $|1,2,1\rangle$, and $|1,1,2\rangle$ are eigenstates of $\hat{H}$ but also of $\hat{K}_x$, $\hat{K}_y$, and $\hat{K}_z$. Is this a coincidence? No. Except for accidental degeneracies, we always find that the states of an observable (H) that are degenerate are so because they are also eigenstates of other observables (K_x, K_y, K_z).

state	K_x	K_y	K_z	E
$\|2,1,1\rangle$	$\frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2$	$\frac{\hbar^2}{2m}\left(\frac{1\pi}{L}\right)^2$	$\frac{\hbar^2}{2m}\left(\frac{1\pi}{L}\right)^2$	$\frac{6\hbar^2\pi^2}{2mL^2}$
$\|1,2,1\rangle$	$\frac{\hbar^2}{2m}\left(\frac{1\pi}{L}\right)^2$	$\frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2$	$\frac{\hbar^2}{2m}\left(\frac{1\pi}{L}\right)^2$	$\frac{6\hbar^2\pi^2}{2mL^2}$
$\|1,1,2\rangle$	$\frac{\hbar^2}{2m}\left(\frac{1\pi}{L}\right)^2$	$\frac{\hbar^2}{2m}\left(\frac{1\pi}{L}\right)^2$	$\frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2$	$\frac{6\hbar^2\pi^2}{2mL^2}$

Table 1: Kinetic and total energy

Finally, note that $\hat{K}_x$, $\hat{K}_y$, and $\hat{K}_z$ all commute with $\hat{H}$. Is this a coincidence? No. We will prove in a later chapter that operators that commute have common eigenfunctions.

Exercise 6 Take $L_x = L_y = L$ and $L_z \neq L$. Does the system have degenerate states? Are these states eigenstates of $\hat{K}_x$, $\hat{K}_y$, and $\hat{K}_z$? How about the case $L_z \neq L_y \neq L_z$?

---

§ 16 *The hydrogen atom.* The hydrogen atom is another system whose energy eigenstates are degenerate. Its energy is (see Metiu, *Quantum Mechanics*, page 300)

$$E_n = -\frac{1}{n^2}\,\frac{\mu e^4}{2(4\pi\epsilon_0)^2\hbar^2} \tag{81}$$

and the states $|n,\ell,m\rangle$ are labeled by three quantum numbers: n controls the energy, ℓ controls the square of the angular momentum (essentially, this is the rotational energy), and m controls the projection of the angular momentum

on the OZ axis. The states $|n, \ell, m\rangle$ are eigenvectors of *three operators*. These operators are (1) the Hamiltonian $\hat{H}$, for which

$$\hat{H}|n, \ell, m\rangle = E_n|n, \ell, m\rangle, \tag{82}$$

(2) the angular momentum squared $\hat{L}^2$, for which

$$\hat{L}^2|n, \ell, m\rangle = \hbar^2\ell(\ell + 1)|n, \ell, m\rangle, \tag{83}$$

and (3) the projection $\hat{L}_z$ of the angular momentum on the OZ axis, for which

$$\hat{L}_z|n, \ell, m\rangle = \hbar m|n, \ell, m\rangle, \tag{84}$$

The states $|2, 0, 0\rangle$, $|2, 1, -1\rangle$, $|2, 1, 0\rangle$, and $|2, 1, 1\rangle$ (remember that ℓ takes only the values $0, 1, \ldots, n-1$ and m takes only the values $-\ell, -\ell+1, \ldots, \ell$) all have the energy

$$E_2 = -\frac{1}{4}\frac{\mu e^4}{2(4\pi\epsilon_0)^2\hbar^2} \tag{85}$$

They have the same energy, but they are different states. The state $|2, 0, 0\rangle$ has no rotational energy. The states $|2, 1, -1\rangle$, $|2, 1, 0\rangle$, and $|2, 1, 1\rangle$ have *the same rotational energy* but differ through the *orientation* of the angular-momentum vector. Are there experiments that distinguish these states? Yes. A hydrogen atom in state $|2, 1, -1\rangle$ emits photons of the same frequency as one in the state $|2, 1, 0\rangle$ but in a different direction.

We have the same situation as for the particle in a cubic box. The degenerate states of the Hamiltonian have the same energy but differ through other properties. The degenerate states are eigenstates of $\hat{H}$ but also of other operators (here $\hat{L}^2$ and $\hat{L}_z$ play roles similar to $\hat{K}_x$, $\hat{K}_y$, and $\hat{K}_z$).

There is also a connection to symmetry: the degeneracy appears because the system is spherically symmetric (the Coulomb attraction $V(r)$ is). Had the interaction been of a form $V(x, y, z)$ with no spherical symmetry, the system would not have had degenerate states.

Section 9.5. Degenerate Eigenvalues: A Numerical Example

§ 17 *A three-dimensional example.* I hope you have accepted that degeneracy is not a freakish accident and it will be good to know why and when a matrix has degenerate eigenstates. Let us start with an example. `Mathematica` tells me (see `WorkBook 9`, Section 3, Cell 1) that the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{86}$$

has the eigenvalues

$$\lambda_1 = 3 \tag{87}$$

$$\lambda_2 = 0 \tag{88}$$

$$\lambda_3 = 0 \tag{89}$$

with corresponding eigenvectors

$$v(1) = \{1, 1, 1\} \tag{90}$$

$$v(2) = \{-1, 0, 1\} \tag{91}$$

$$v(3) = \{-1, 1, 0\} \tag{92}$$

I calculated $\langle v(i) \,|\, v(j) \rangle$ and found

$$\langle v(1) \,|\, v(1) \rangle = 3, \quad \langle v(2) \,|\, v(2) \rangle = 2, \quad \langle v(3) \,|\, v(3) \rangle = 2 \tag{93}$$

The eigenvectors are *not normalized*. I also calculated that

$$\langle v(1) \,|\, v(2) \rangle = 0, \quad \langle v(1) \,|\, v(3) \rangle = 0, \quad \langle v(2) \,|\, v(3) \rangle = 1 \tag{94}$$

This means that the degenerate eigenvectors $v(2)$ and $v(3)$ are *not orthogonal to each other* but they are orthogonal to the eigenvector corresponding to a different eigenvalue (namely to $v(1)$).

In quantum mechanics we use the eigenvectors to calculate probabilities, or the mean value of an observable, or as an orthonormal basis set to represent other vectors and/or operators. We would like it, if at all possible, to convert the set of eigenvectors *into another set of eigenvectors that are orthonormal and correspond to the same eigenvalues*. I will show below that this is possible. First we start with the set of degenerate eigenvectors and use the Gram-Schmidt procedure to form a new set of vectors that are orthogonal to each other. Then we normalize them. The vectors produced by this procedure are still degenerate eigenvectors corresponding to the eigenvalue 0, and are still orthogonal to the either eigenvectors ($v(1)$ in this example). This follows from the fact that the vectors generated by Gram-Schmidt are linear combinations of the degenerate eigenvectors.

§ 18 *Completeness.* If some of the eigenvectors are degenerate do they still provide a complete basis set? The answer is yes. We know that because we have just argued that we can convert them into a set of orthonormal eigenvectors; any orthonormal set of three vectors is a complete basis set.

We can approach the same question from a different direction. When we ask whether the set $\{v(1), v(2), v(3)\}$ is complete we are asking whether any vector x in the space can be written in the form

$$x = \alpha_1 v(1) + \alpha_2 v(2) + \alpha_3 v(3) \tag{95}$$

where α_1, α_2, α_3 are suitably chosen numbers.

The answer will be yes if the three vectors are *linearly independent*, which means that there are *no numbers* β_1, β_2, β_3 such that

$$\beta_1 v(1) + \beta_2 \vec{v}(2) + \beta_3 \vec{v}(3) = 0 \tag{96}$$

(the selection $\beta_1 = \beta_2 = \beta_3 = 0$ does not count). Another way to put this is that none of $v(1)$, $v(2)$, $v(3)$ can be written as a linear combination of the others. How do we establish that this is true? Calculating the left-hand side of Eq. 96 for all possible numbers would take too long. Fortunately, there is a short-cut we can use, based on the theory of determinants. Form the matrix

$$B = \begin{pmatrix} v_1(1) & v_2(1) & v_3(1) \\ v_1(2) & v_2(2) & v_3(2) \\ v_1(3) & v_2(3) & v_3(1) \end{pmatrix} \tag{97}$$

where $v_i(\alpha)$ is the i-th component of the vector $v(\alpha)$.

If a row can be written as a linear combination of the other rows, then the determinant of B is zero (and vice versa). Therefore

$$\det B \neq 0 \tag{98}$$

ensures that the three vectors are linearly independent. Straightforward calculation (or `Mathematica`) tells me that in this case

$$\det B = -3 \tag{99}$$

We conclude that the three eigenvectors *are linearly independent*. Therefore the three eigenvectors form a complete basis set, but they are not orthonormal.

Exercise 7 Suppose three 3-dimensional vectors lie in the same plane. Are they linearly independent? What about three vectors for which two are in the same plane but not parallel and the third is perpendicular to that plane?

---

§ 19 *Orthogonalization.* We can use the Gram-Schmidt orthogonalization procedure (see Chapter 3) to convert the degenerate eigenvectors $v(2)$ and $v(3)$ into two orthonormal vectors $x(2)$ and $x(3)$. These are

$$x(2) = v(2) \tag{100}$$

$$x(3) = v(3) - \frac{v(2)}{\langle x(2) \,|\, x(2)\rangle}\langle x(2) \,|\, v(3)\rangle \tag{101}$$

The calculation was performed in Section 2, Cell 2 of `WorkBook9` and the result is

$$x(2) = v(2) = \{-1, 0, 1\} \tag{102}$$

$$x(3) = \{-\frac{1}{2}, 1, -\frac{1}{2}\} \tag{103}$$

It is easy to verify that $Mx(2) = 0 = 0 \times x(2)$ and $Mx(3) = 0 = 0 \times x(3)$, which means that $x(2)$ and $x(3)$ are eigenstates of M corresponding to the eigenvalue equal to zero. The Gram-Schmidt procedure converts degenerate, non-orthogonal eigenvectors belonging to the eigenvalue 0 into orthogonal eigenvectors belonging to the same eigenvalue.

We can verify that $\langle v(1)\,|\,x(2)\rangle = \langle v(1)\,|\,x(3)\rangle = \langle x(2)\,|\,x(3)\rangle = 0$; the vectors $x(i)$ are orthogonal, as expected.

Next we normalize $x(1)$, $x(2)$, and $x(3)$, by using

$$s(1) = \frac{x(1)}{\sqrt{\langle x(1)\,|\,x(1)\rangle}} \tag{104}$$

$$s(2) = \frac{x(2)}{\sqrt{\langle x(2)\,|\,x(2)\rangle}} \tag{105}$$

$$s(3) = \frac{x(3)}{\sqrt{\langle x(3)\,|\,x(3)\rangle}} \tag{106}$$

The orthonormal vectors are therefore (see `WorkBook9`)

$$s(1) = \left\{ \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right\} \tag{107}$$

$$s(2) = \left\{ -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right\} \tag{108}$$

$$s(3) = \left\{ -\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}, -\frac{1}{\sqrt{6}} \right\} \tag{109}$$

Exercise 8 Verify that (1) $Ms(i) = \lambda_i s(i)$, $i = 1, 2, 3$, with $\lambda_1 = 3$, $\lambda_2 = 0$, $\lambda_3 = 0$; and (2) $\langle s(i)\,|\,s(j)\rangle = \delta_{ij}$, $i, j = 1, 2, 3$.

What you have seen in this example is general.

1. The eigenvectors of a Hermitian matrix are linearly independent.

2. The eigenvectors corresponding to different eigenvalues are automatically orthogonal.

3. The degenerate eigenvectors are not necessarily orthogonal to each other. One can always convert them (by Gram-Schmidt) into a set of degenerate eigenvectors that are orthogonal to each other and to all other eigenvectors.

4. Any nonzero vector can be normalized by dividing it by its norm ($x/\sqrt{x \cdot x}$ is normalized).

§ 20 *Find the eigenvalues of the matrix M (Eq. 86) "by hand".* You have learned enough to be able to use the computer and generate orthonormal eigenvectors. It would not hurt, however, to understand how degeneracy appears, by doing an eigenvalue calculation, step by step, to look at the details.

The eigenvalue problem for matrix M is

$$Mx = \lambda x \tag{110}$$

For the matrix given by Eq. 86, this is the same as the system of equations

$$(1-\lambda)x_1 + x_2 + x_3 = 0 \tag{111}$$

$$x_1 + (1-\lambda)x_2 + x_3 = 0 \tag{112}$$

$$x_1 + x_2 + (1-\lambda)x_3 = 0 \tag{113}$$

where x_1, x_2, x_3 are the components of the vector x.

The characteristic polynomial is

$$\det[M - \lambda I] = \det \begin{pmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{pmatrix} = 3\lambda^2 - \lambda^3 \tag{114}$$

The eigenvalues are the roots of this polynomial, i.e. the solutions of the equation

$$3\lambda^2 - \lambda^3 = 0 \tag{115}$$

They are (see `WorkBook9`, Section 3, Cell 4)

$$\lambda_1 = 3 \tag{116}$$

$$\lambda_2 = 0 \tag{117}$$

$$\lambda_3 = 0 \tag{118}$$

We can calculate the eigenvector $x(1)$ corresponding to $\lambda_1 = 3$ as we did before (this eigenvector is not degenerate). The equation $(M - \lambda_1 I)\vec{x}(1) = \mathbf{0}$ is shorthand for the system of linear equations:

$$(1-3)x_1(1) + x_2(1) + x_3(1) = 0 \tag{119}$$

$$x_1(1) + (1-3)x_2(1) + x_3(1) = 0 \tag{120}$$

$$x_1(1) + x_2(1) + (1-3)x_3(1) = 0 \tag{121}$$

These equations have a solution different from $x_1(1) = x_2(1) = x_3(1) = 0$ because we use $\lambda = 3$, which makes the determinant in Eq. 114 equal to zero. To find this solution, we take two equations and solve for $x_2(1)$ and $x_3(1)$. The result will depend on $x_1(1)$, which is left as a parameter to be determined later.

I solved Eqs. 119 and 120 and obtained

$$x_2(1) = x_1(1) \tag{122}$$

$$x_3(1) = x_1(1) \tag{123}$$

The eigenvector $x(1)$ corresponding to the eigenvalue $\lambda_1 = 3$ is therefore

$$x(1) = \{x_1(1), x_1(1), x_1(1)\} = x_1(1)\{1, 1, 1\} \tag{124}$$

This is a perfectly fine result. The presence of $x_1(1)$ in it does not bother me since it can be determined by forcing $x(1)$ to be normalized. The normalized eigenvector is

$$s(1) = \frac{x(1)}{\sqrt{\langle x(1) \,|\, x(1)\rangle}} = \left\{\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right\} \tag{125}$$

Notice that $x_1(1)$ disappears from this expression. The non-degenerate eigenvector was calculated as in Section 9.2. The fact that the other two eigenvectors are degenerate does not affect this calculation.

The situation is different for the degenerate eigenvectors $x(2)$ and $x(3)$. They satisfy

$$(M - 0\,I)x(2) = 0 \tag{126}$$

and

$$(M - 0\,I)x(3) = 0 \tag{127}$$

because the eigenvalues λ_2 and λ_3 are equal to zero. Note that *these two equations are identical*. However, we must have $x(2) \neq x(3)$ because they are different eigenstates.

We seem to have reached an impasse but we should not lose faith in the general theory. Let us ignore the complication and proceed. To obtain the eigenvectors for the zero eigenvalue, we make $\lambda_2 = 0$ in Eqs. 111–113. This gives

$$x_1(2) + x_2(2) + x_3(2) \;\; = \;\; 0 \tag{128}$$

$$x_1(2) + x_2(2) + x_3(2) = 0 \tag{129}$$

$$x_1(2) + x_2(2) + x_3(2) = 0 \tag{130}$$

We might have expected a system of three *different* equations but we get instead *three identical equations.* It turns out that this is what is needed in order to obtain distinct, degenerate eigenvectors.

All I can conclude from Eqs. 128–130 is that

$$x_1(2) = -x_2(2) - x_3(2) \tag{131}$$

Therefore, the eigenvectors corresponding to $\lambda_2 = 0$ and $\lambda_3 = 0$ are both of the form

$$x(\alpha) = \{-x_2(\alpha) - x_3(\alpha),\ x_2(\alpha),\ x_3(\alpha)\}, \quad \alpha = 2, 3 \tag{132}$$

This is all we can get out of the eigenvalue equation which does not determine x_2 and x_3; *any vector of this form satisfies the equation* $Mx = 0\,x$, as you can easily verify. In addition, any vector of the form in Eq. 132 is orthogonal to $s(1)$ because $s(1)$ corresponds to a different eigenvalue of M. In `Workbook9`, I verified that this is indeed the case.

Exercise 9 For $x = \{-a - b,\ a,\ b\}$, test that $Mx = 0\,x$ and $\langle x \,|\, s(1)\rangle = 0$.

To get two non-zero eigenvectors corresponding to the eigenvalue $\lambda = 0$, I first pick *arbitrarily* $x_2(2) = 1$ and $x_3(2) = 2$, and obtain (from the general form Eq. 132)

$$x(2) = \{-3, 1, 2\} \tag{133}$$

Then I pick a second vector by choosing arbitrarily[3] $x_2(3) = 0$ and $x_3(3) = 1$ and obtain

$$x(3) = \{-1, 0, 1\} \tag{134}$$

Physics requires that these two vectors be normalized and orthogonal. The Gram-Schmidt procedure will make them orthogonal, and the orthogonalized eigenvectors are

$$v(2) = x(2) = \{-3, 1, 2\} \tag{135}$$

$$v(3) = x(3) - \frac{v(2)}{\langle v(2) \,|\, v(2)\rangle}\langle v(2) \,|\, x(3)\rangle = \left\{\frac{1}{14}, -\frac{5}{14}, \frac{2}{7}\right\} \tag{136}$$

Exercise 10 Show that $\langle \vec{s}(1) \,|\, \vec{v}(2)\rangle = \langle \vec{s}(1) \,|\, \vec{v}(3)\rangle = \langle \vec{v}(2) \,|\, \vec{v}(3)\rangle = 0$ and $M\vec{v}(i) = 0\vec{v}(i)$ for $i = 2, 3$.

The vectors $v(2)$ and $v(3)$ are not normalized. The respective normalized eigenvectors are

$$s(2) = \frac{v(2)}{\sqrt{\langle v(2) \,|\, v(2)\rangle}} = \left\{-\frac{3}{\sqrt{14}}, \frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}\right\} \tag{137}$$

$$s(3) = \frac{v(3)}{\sqrt{\langle v(3) \,|\, v(3)\rangle}} = \left\{\frac{1}{\sqrt{42}}, -\frac{5}{\sqrt{42}}, \frac{2\sqrt{2}}{\sqrt{21}}\right\} \tag{138}$$

These vectors $s(2)$ and $s(3)$ are degenerate eigenvectors corresponding to the eigenvalues $\lambda_2 = \lambda_3 = 0$. The system $\{s(1), s(2), s(3)\}$ is orthonormal.

Note that we picked arbitrary values for $x_2(2)$ and $x_3(2)$. Therefore what I have proven is that any vector of the form

$$x = \{-a - b,\ a,\ b\}$$

[3] or almost so: make sure not to choose a multiple of the first vector

is an eigenvector corresponding to the degenerate eigenvalue $\lambda = 0$, regardless of the values of a and b.

How does the degenerate problem differ from the non-degenerate one? For a non-degenerate eigenvalue, we lose an equation and we can solve the remaining ones. The eigenvector given by them has an undetermined component (since we lost an equation) and we find it by imposing normalization. If an eigenvalue is doubly degenerate, *we lose two equations*. The eigenvectors have two undetermined components and we determine them by forcing the degenerate eigenvectors to be orthonormal. If you have a 9×9 matrix and an eigenvalue is six-fold degenerate, you lose seven equations; and so forth.