Chapter 7. The Eigenvalue Problem
Section 7.1. The general case

§ 1 Introduction. The eigenvalue equation of an operator \( \hat{O} \) is

\[ \hat{O}|x\rangle = \lambda|x\rangle \]  

where \( \lambda \) is an unknown number and \( |x\rangle \) is an unknown ket. This equation has many solutions \( |x_\alpha\rangle \), which are called eigenstates, or eigenkets, or eigenvectors of \( \hat{O} \). For each solution \( |x_\alpha\rangle \), there is a number \( \lambda_\alpha \), called the eigenvalue corresponding to the eigenket \( |x_\alpha\rangle \).

I remind you that if \( \hat{A} \) represents an observable then the eigenstates of \( \hat{A} \) are the pure states of that observable. In addition, the spectrum of \( A \) (the values that the observable \( A \) can take) consists of the eigenvalues of \( \hat{A} \).

Solving the eigenvalue equation for the operators representing observables is a matter of greatest importance to quantum mechanics. In most cases encountered in practice, you will know the operator and will want to calculate some of the eigenvalues and eigenkets.

If one works with the Schrödinger representation (which is the coordinate representation), the operators of interest are differential operators acting on wave functions that are elements of \( L^2 \). For example, the energy operator for a harmonic oscillator is

\[ \hat{H} = -\frac{\hbar^2}{2\mu} \frac{d^2}{dx^2} - kx^2 \]  

and the eigenvalue equation is

\[ \hat{H}\psi(x) \equiv -\frac{\hbar^2}{2\mu} \frac{d^2\psi(x)}{dx^2} - kx^2\psi(x) = E\psi(x) \]  

The eigenvalue equation for the energy of the harmonic oscillator has an analytic solution (see Metiu, Quantum Mechanics) and so do those equations for the particle in a box and for the hydrogen atom. This situation is not general.

One of the most frequently used methods for solving such equations numerically is to turn them into a matrix equation. This is why we dedicate a chapter to studying the eigenvalue problem for matrices. In practice, we use computers and software provided by experts to solve the eigenvalue problem for a given matrix. 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. For demanding calculations, we must use Fortran, C, or C++, since Mathematica is less efficient. These computer languages have libraries of programs designed to give you the eigenvalues and the eigenvectors when you give them the matrix.

§ 2 Some general considerations. It turns out that the eigenvalue problem $A|\psi\rangle = \lambda|\psi\rangle$ has many solutions and not all of them are physically meaningful. We are interested in this eigenvalue problem because, if it arises from an observable, the eigenvalues are the spectrum of the observable and some of the eigenfunctions are the pure states of the observable. In addition, the operator representing the observable is Hermitian, because its spectrum, which consists of all the values the observable can take, just contain only real numbers.

Before we try to solve the eigenvalue problem, we need to decide how we recognize which eigenstates correspond to pure states: these are the only eigenstates of interest to us.

In Chapter 2, we examined the definition of the probability and concluded that kets $|a_i\rangle$ and $|a_j\rangle$ representing pure states must satisfy the orthonormalization condition

$$\langle a_i | a_j \rangle = \delta_{ij} \quad (4)$$

§ 3 Normalization. Let us assume that we are interested in solving the eigenvalue equation

$$\hat{A}|\psi\rangle = \lambda|\psi\rangle \quad (5)$$

It is easy to see that if $\alpha$ is a complex number and $|\psi\rangle$ satisfies Eq. 5 then $|\eta\rangle = \alpha|\psi\rangle$ satisfies the same equation. For every given eigenvalue we have as many eigenstates as complex numbers. The one that is a pure state must be normalized. We can therefore determine the value of $\alpha$ by requiring that $|\eta\rangle$ satisfies

$$\langle \eta | \eta \rangle = 1 \quad (6)$$

Using Eq. 5 in Eq. 6 (and the properties of the scalar product) leads to

$$\alpha^*\alpha \langle \psi | \psi \rangle = 1 \quad (7)$$

or

$$\alpha^*\alpha = \frac{1}{\langle \psi | \psi \rangle} \quad (8)$$
This gives

\[ \alpha = \frac{e^{i\phi}}{\sqrt{\langle \psi | \psi \rangle}} \]

(9)

where \( \phi \) is a real number, which cannot be determined by solving Eq. 8. The ket

\[ |\eta \rangle = \frac{e^{i\phi}}{\sqrt{\langle \psi | \psi \rangle}} |\psi \rangle \]

(10)

is normalized and therefore is a pure state of \( \hat{A} \) corresponding to the eigenvalue \( \lambda \). If the system is in the state \( |\eta \rangle \) and we measure \( A \), we are certain that the result is \( \lambda \).

In practice, most computer programs that calculate eigenstates do not normalize them. If the computer produces \( |\psi \rangle \) as the eigenket corresponding to eigenvalue \( \lambda \), you must use Eq. 10 to find the eigenstate \( |\eta \rangle \) that is a pure state. Since \( e^{i\phi} \) in Eq. 10 is an irrelevant phase factor, we can drop it (i.e. take \( \phi = 0 \)); the pure state corresponding to \( \lambda \) is

\[ |\eta \rangle = \frac{|\psi \rangle}{\sqrt{\langle \psi | \psi \rangle}} \]

(11)

§ 4 Orthogonality. The pure states must also be orthogonal. Because an operator \( \hat{A} \), representing an observable \( A \), is always Hermitian, one can prove that eigenstates corresponding to different eigenvalues must be orthogonal. That is, if

\[ \hat{A}|\eta_i \rangle = \lambda_i |\eta_i \rangle \quad \text{and} \quad \hat{A}|\eta_j \rangle = \lambda_j |\eta_j \rangle \]

(12)

and

\[ \lambda_i \neq \lambda_j \]

(13)

then

\[ \langle \eta_i | \eta_j \rangle = 0 \]

(14)

The proof is easy. We have

\[ \langle \hat{A} | \eta_i \rangle - \langle \eta_i | \hat{A} \rangle | \eta_j \rangle = 0 \]

(15)

because \( \hat{A} \) is Hermitian. Using Eqs. 12 in Eq. 15 gives

\[ (\lambda_i - \lambda_j) \langle \eta_i | \eta_j \rangle = 0 \]
When \( \lambda_i \neq \lambda_j \), we must have

\[
\langle \eta_i | \eta_j \rangle = 0,
\]

which is what we wanted to prove.

So, when you solve an eigenvalue equation for an observable, the eigenstates corresponding to different eigenvalues must be orthogonal. If they are not, there is something wrong with the program calculating the eigenvectors.

What happens if some eigenvalues are equal to each other? Physicists call such eigenvalues ‘degenerate’, mathematicians call them ‘multiple’. If \( |\eta_i\rangle \) and \( |\eta_j\rangle \) belong to the same eigenvalue (i.e. if \( A|\eta_i\rangle = \lambda|\eta_i\rangle \) and \( A|\eta_j\rangle = \lambda|\eta_j\rangle \)) then \( |\eta_i\rangle \) and \( |\eta_j\rangle \) are under no obligation to be orthogonal to each other (only to, according to the fact in the previous §, all eigenvectors belonging to eigenvalues that differ from \( \lambda \)). Most numerical procedures for solving eigenvalue problems will provide non-orthogonal degenerate eigenvectors. These can be orthogonalized with the Gram-Schmidt procedure and then normalized. The resulting eigenvectors are now pure states. You will see how this is done in an example given later in this chapter.

It is important to keep in mind this distinction between eigenstates and pure states. All pure states are eigenstates but not all eigenstates are pure states.

§ 5 The eigenvalue problem in matrix representation: a review. Let \( |a_1\rangle, |a_2\rangle, \ldots, |a_n\rangle, \ldots \) be a complete orthonormal basis set. This set need not be a set of pure states of an observable; we can use any complete basis set that a mathematician can construct. Orthonormality means that

\[
\langle a_i | a_j \rangle = \delta_{ij} \text{ for all } i, j
\]

Because the set is complete, we have

\[
\hat{I} \sim \sum_{n=1}^{N} |a_n\rangle \langle a_n|\]

Eq. 17 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 to turn it into the following \( N \) equations:

\[
\langle a_m | \hat{A} | \psi \rangle = \lambda \langle a_m | \psi \rangle, \quad m = 1, 2, \ldots, N
\]
Now insert $\hat{I}$, as given by Eq. 17, between $\hat{A}$ and $|\psi\rangle$ to obtain

$$\sum_{n=1}^{N} \langle a_m | \hat{A} | a_n \rangle \langle a_n | \psi \rangle = \lambda \langle a_m | \psi \rangle, \quad m = 1, 2, \ldots, N$$  \hspace{1cm} (19)

The complex numbers $\langle a_n | \psi \rangle$ are the coordinates of $|\psi\rangle$ in the $\{|a_n\rangle\}_{n=1}^{N}$ representation. If we know them, we can write $|\psi\rangle$ as

$$|\psi\rangle = \sum_{n=1}^{N} |a_n\rangle \langle a_n | \psi \rangle$$  \hspace{1cm} (20)

It is easy to rewrite Eq. 19 in a form familiar from linear matrix algebra. Let us denote

$$\psi_i \equiv \langle a_i | \psi \rangle, \quad i = 1, 2, \ldots, N$$  \hspace{1cm} (21)

and

$$A_{mn} \equiv \langle a_m | \hat{A} | a_n \rangle$$  \hspace{1cm} (22)

With this notation, Eq. 19 becomes

$$\sum_{n=1}^{N} A_{mn} \psi_n = \lambda \psi_m, \quad m = 1, 2, \ldots, N$$  \hspace{1cm} (23)

and Eq. 20,

$$|\psi\rangle = \sum_{n=1}^{N} |a_n\rangle \psi_n$$  \hspace{1cm} (24)

The sum in Eq. 23 is the rule by which the matrix $A$, having the elements $A_{mn}$, acts on the vector $\psi$, having the coordinates $\psi_n$. This equation is called the eigenvalue problem for the matrix $A$ and it is often written as

$$A \psi = \lambda \psi$$  \hspace{1cm} (25)

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

$$\begin{pmatrix}
A_{11} & A_{12} & \cdots & A_{1N} \\
A_{21} & A_{22} & \cdots & A_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
A_{N1} & A_{N2} & \cdots & A_{NN}
\end{pmatrix} \begin{pmatrix}
\psi_1 \\
\psi_2 \\
\vdots \\
\psi_N
\end{pmatrix} = \lambda \begin{pmatrix}
\psi_1 \\
\psi_2 \\
\vdots \\
\psi_N
\end{pmatrix}$$  \hspace{1cm} (26)

Eqs. 25 and 26 are different ways of writing Eq. 23, which, in turn, is the representation of the equation $\hat{A}|\psi\rangle = \lambda |\psi\rangle$ in the basis set $\{|a_1\rangle, \ldots, |a_N\rangle\}$. 
Normally in quantum mechanics we choose the basis set \{\ket{a_1}, \ldots, \ket{a_n}\} 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 \(\ket{\psi}\) or \(\lambda\), and our mission is to find them from \(\hat{A}\ket{\psi} = \lambda \ket{\psi}\).

We have converted this operator equation, by using the orthonormal basis set \{\ket{a_1}, \ldots, \ket{a_N}\}, into the matrix equation Eq. 23 (or Eq. 25 or Eq. 26). In Eq. 23, we know \(A_{mn}\) but we do not know \(\psi = \{\psi_1, \psi_2, \ldots, \psi_N\}\) or \(\lambda\). We must calculate them from Eq. 23. Once we know \(\psi_1, \psi_2, \ldots, \psi_N\), we can calculate \(\ket{\psi}\) from

\[
\ket{\psi} = \sum_{n=1}^{N} \ket{a_n} \langle a_n | \psi \rangle \equiv \sum_{n=1}^{N} \psi_n \ket{a_n}
\] (27)

§ 6 Use a computer. Calculating eigenvalues and eigenvectors of a matrix “by hand” is extremely tedious, especially since the dimension \(N\) is often very large. Most computer languages (including Fortran, C, C++, Mathematica, Maple, Mathcad, Basic) have libraries of functions that given a matrix will return its eigenvalues and eigenvectors.

I will not discuss the numerical algorithms used for finding eigenvalues. Quantum mechanics could leave that to experts without suffering much harm. The Mathematica file Linear algebra for quantum mechanics.nb shows how to use Mathematica to perform calculations with vectors and matrices.

Nevertheless, it is important to understand some of the theory related to the eigenvalue problem, since much of quantum mechanics relies on it. I present some of it in what follows.

§ 7 The eigenvalue problem and systems of linear equations. The eigenvalue problem Eq. 23 can be written as

\[
\sum_{n=1}^{N} (A_{mn} - \lambda \delta_{mn}) \psi_n = 0, \quad m = 1, 2, \ldots, N
\] (28)

or

\[
\begin{pmatrix}
A_{11} - \lambda & A_{12} & A_{13} & \cdots & A_{1N} \\
A_{21} & A_{22} - \lambda & A_{23} & \cdots & A_{2N} \\
A_{31} & A_{32} & A_{33} - \lambda & \cdots & A_{3N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A_{N1} & A_{N2} & A_{N3} & \cdots & A_{NN} - \lambda
\end{pmatrix}
\begin{pmatrix}
\psi_1 \\
\psi_2 \\
\psi_3 \\
\vdots \\
\psi_N
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
\]

(29)
The column in the right-hand side is the zero vector. We can regard \( \delta_{mn} \) as being the elements of a matrix \( I \), which is called the unit matrix (or the identity matrix):

\[
I = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix}
\]

(30)

This has the property

\[
(Iv)_m = \sum_{n=1}^{N} I_{mn}v_n = \sum_{n=1}^{N} \delta_{mn}v_n = v_m
\]

The \( m \)-th element of the vector \( Iv \) is \( v_m \); this means that

\[
Iv = v
\]

(31)

The definition of \( I \) allows us to write Eq. 29 as

\[
(A - \lambda I)\psi = 0
\]

(32)

A system of equations of the form

\[
Mv = 0,
\]

(33)

where \( M \) is an \( N \times N \) matrix and \( v \) is an \( N \)-dimensional vector, is called a homogeneous system of \( N \) linear equations. The name ‘homogeneous’ is given because the right-hand side is zero; this distinguishes it from

\[
Mv = u
\]

(34)

which is an inhomogeneous system of \( N \) linear equations.

Example. Let

\[
A = \begin{pmatrix}
3 & 2 & 4 \\
2 & 1.2 & 3.1 \\
4 & 3.1 & 4
\end{pmatrix}
\]

(35)
This is a Hermitian matrix with real elements. Let \( \psi = \{\psi_1, \psi_2, \psi_3\} \) be a vector. Then \( A\psi = 0 \) means (use \( (A\psi)_i = \sum_{j=1}^{3} A_{ij}\psi_j \))

\[
A\psi = \begin{pmatrix}
A_{11}\psi_1 + A_{12}\psi_2 + A_{13}\psi_3 \\
A_{21}\psi_1 + A_{22}\psi_2 + A_{23}\psi_3 \\
A_{31}\psi_1 + A_{32}\psi_2 + A_{33}\psi_3 \\
\end{pmatrix} = \begin{pmatrix}
3\psi_1 + 2\psi_2 + 4\psi_3 \\
2\psi_1 + 1.2\psi_2 + 3.1\psi_3 \\
4\psi_1 + 3.1\psi_2 + 4\psi_3 \\
\end{pmatrix} = 0 \tag{36}
\]

If a vector is equal to zero then \textit{all its components} are zero, and Eq. 36 means

\[
\begin{align*}
3\psi_1 + 2\psi_2 + 4\psi_3 &= 0 \tag{37} \\
2\psi_1 + 1.2\psi_2 + 3.1\psi_3 &= 0 \tag{38} \\
4\psi_1 + 3.1\psi_2 + 4\psi_3 &= 0 \tag{39}
\end{align*}
\]

The notation \( A\psi = 0 \) is a shorthand for these three equations. This is a homogeneous system of three linear equations.

Similarly if \( \phi = \{2.1, 6.2, 4.3\} \) then

\[
A\psi = \phi \tag{40}
\]

is shorthand for

\[
\begin{align*}
3\psi_1 + 2\psi_2 + 4\psi_3 &= 2.1 \tag{41} \\
2\psi_1 + 1.2\psi_2 + 3.1\psi_3 &= 6.2 \tag{42} \\
4\psi_1 + 3.1\psi_2 + 4\psi_3 &= 4.3 \tag{43}
\end{align*}
\]

This is an inhomogeneous system of three linear equations. \textit{If the matrix A has an inverse} \( A^{-1} \), we can solve Eq. 40 by using

\[
A^{-1}A\psi = A^{-1}\phi \tag{44}
\]

and

\[
A^{-1}A\psi = I\psi = \psi \tag{45}
\]

which together lead to

\[
\psi = A^{-1}\phi \tag{46}
\]

You have learned in your mathematics courses about Cramer’s rule, which allows you to solve Eq. 40. This is tedious to implement and I prefer to
use Mathematica to give me the inverse matrix (see the file \texttt{WorkBook7 The eigenvalue problem.nb}). Mathematica tells me that

$$A^{-1} = \begin{pmatrix} 160.333 & -146.667 & -46.6667 \\ -146.667 & 133.333 & 43.3333 \\ -46.6667 & 43.3333 & 13.3333 \end{pmatrix}$$  

(47)

Acting with $A^{-1}$ on $\phi$ gives (see \texttt{WorkBook7.nb})

$$\psi = A^{-1}\phi = \{-773.3, 705.0, 228.0\}$$  

(48)

\textit{End of Example}

\textbf{Exercise 1} Prove that if a matrix $H$ is Hermitian then so is its inverse.

Note a very important fact. If $A^{-1}$ exists then $A\psi = \phi$ has a solution. This also means that the solution of $A\psi = 0$ is $\psi = 0$ (because $A^{-1}A\psi = A^{-1}0 = 0$ and $A^{-1}A\psi = I\psi = \psi$). The solution $\psi = 0$ is of no interest, and we say that if $A^{-1}$ exists then $A\psi = 0$ does not have a solution.

How does this relate to the eigenvalue problem? The eigenvalue equation written in the form

$$(A - \lambda I)\psi = 0$$  

(49)

is a homogenous system of equations. If the operator

$$(\hat{A} - \lambda \hat{I})^{-1}$$  

(50)

exists then $(A - \lambda I)\psi = 0$ does not have a solution.

Therefore $(A - \lambda I)\psi = 0$ has a solution if and only if $(A - \lambda I)^{-1}$ does not exist, which happens exactly when

$$\text{det}(A - \lambda I) = 0$$  

(51)

Note that this condition does not involve the unknown vector $\psi$; it depends only on $\lambda$ (which we don’t know) and on the matrix elements $A_{ij}$ (which we do know). This is an equation for $\lambda$. One can prove, based on the properties of determinants, that the left-hand side of Eq. 51 is a polynomial of order $N$ in $\lambda$, which is called the \textit{characteristic polynomial} of the matrix $A$. The fundamental theorem of algebra tells us that Eq. 51 must therefore have
7. The Eigenvalue Problem, December 17, 2009

There are $N$ solutions. Because they are eigenvalues of $A$ and $A$ is Hermitian, these solutions must all be real numbers.

I will assume here that you have heard of determinants. If not, take heart: computers can calculate them for you. *Mathematica* has the function `Det[A]` that returns the determinant of the matrix $A$.

§ 8 Back to physics. If $A$ is an observable then the eigenvalues of the matrix $A$ are the values that we observe when we measure the observable $A$. The “quantization” of the values that the observable can take occurs because $(A - \lambda I)\psi = 0$ has solutions only for those special values of $\lambda$ for which the matrix $A - \lambda I$ does not have an inverse. These are the values of $\lambda$ 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. As we increase the number of functions in the orthonormal basis set, we increase the number of eigenvalues. If we use an infinite basis set, we will have an infinite number of eigenvalues. These eigenvalues will be discrete, however.

We cannot possibly use an infinite basis set and therefore we never get the exact eigenvalues and eigenfunctions. By luck, or perhaps because of some mathematical theorem unknown to me, if we are intelligent in our choice of basis set, we get a good approximation to the smallest eigenvalues even with a relatively small basis set. As you’ll see in an example given later, using roughly 40 basis functions gives accurate values for about the 20 lowest eigenvalues.

§ 9 An example of an eigenvalue problem: nondegenerate eigenvalues. Let us examine the eigenvalue problem for the matrix $A$ given by Eq. 35. The characteristic equation is (see Eq. 51)

$$
\begin{vmatrix}
3 - \lambda & 2 & 4 \\
2 & 1.2 - \lambda & 3.1 \\
4 & 3.1 & 4 - \lambda \\
\end{vmatrix} = 0
$$

(52)

If you find tedious algebra soothing, you can calculate this determinant. I prefer to use *Mathematica*, and the result is (see Section 2 of the file `WorkBook7 The eigenvalue problem.nb`)

$$
-\lambda^3 + 8.2\lambda^2 + 9.21\lambda - 0.03 = 0
$$

(53)
This is the characteristic polynomial of the matrix $A$. It is a third-order polynomial because we used a three-dimensional basis set and therefore $A$ is a $3 \times 3$ matrix.

The roots of this equation are the eigenvalues of $A$. They are (see Section 2 of WorkBook7)

$$
\begin{align*}
\lambda_1 &= -1.00391 \\
\lambda_2 &= 0.0032479 \\
\lambda_3 &= 9.20007
\end{align*}
$$

(54)

The eigenvalues are 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 $\lambda_1$, insert $\lambda_1$ in the equation $A\psi = \lambda\psi$ to obtain

$$A\psi(1) - \lambda_1 \psi(1) = 0,
$$

(55)

We have modified the notation so that $\psi(1) = \{\psi_1(1), \psi_2(1), \psi_3(1)\}$ is the eigenvector corresponding to the eigenvalue $\lambda_1$. If I insert $\lambda_2$ in $A\psi = \lambda\psi$, I get a different equation

$$A\psi(2) - \lambda_2 \psi(2) = 0
$$

(56)

for a different eigenvector, $\psi(2)$, corresponding to the eigenvalue $\lambda_2$. Similarly,

$$A\psi(3) - \lambda_3 \psi(3) = 0
$$

(57)

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

Each symbol $\psi(1)$, $\psi(2)$, $\psi(3)$ represents a vector in $\mathbb{R}^3$. The components of $\psi(\alpha)$, $\alpha = 1, 2, 3$, are denoted by

$$\psi(\alpha) = \{\psi_1(\alpha), \psi_2(\alpha), \psi_3(\alpha)\}
$$

(58)

With this notation,

$$
(A - \lambda_1 I)\psi(1) = \begin{pmatrix}
3 - \lambda_1 & 2 & 4 \\
2 & 1.2 - \lambda_1 & 3.1 \\
4 & 3.1 & 4 - \lambda_1
\end{pmatrix}
\begin{pmatrix}
\psi_1(1) \\
\psi_2(1) \\
\psi_3(1)
\end{pmatrix} = 0
$$
This is the system of equations
\[
(3 - \lambda_1)\psi_1(1) + 2\psi_2(1) + 4\psi_3(1) = 0 \quad (59)
\]
\[
2\psi_1(1) + (1.2 - \lambda_1)\psi_2(1) + 3.1\psi_3(1) = 0 \quad (60)
\]
\[
4\psi_1(1) + 3.1\psi_2(1) + (4 - \lambda_1)\psi_3(1) = 0 \quad (61)
\]

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, since, as noted earlier, we can fix \(\beta\) by requiring that the eigenvector be normalized. This is the only meaningful eigenvector in quantum mechanics.

I can look at Eqs. 59–61 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 \(\{0, 0, 0\}\)? Since this is a homogeneous system of equations, the answer is yes: we have chosen \(\lambda = \lambda_1\), which means that \(\text{det}[A - \lambda I] = 0\) and the system has a solution.

If we divide Eqs. 59–61 by \(\psi_3(1)\), we obtain
\[
(3 - \lambda_1)\frac{\psi_1(1)}{\psi_3(1)} + 2\frac{\psi_2(1)}{\psi_3(1)} = -4
\]
\[
2\frac{\psi_1(1)}{\psi_3(1)} + (1.2 - \lambda_1)\frac{\psi_2(1)}{\psi_3(1)} = -3.1
\]
\[
4\frac{\psi_1(1)}{\psi_3(1)} + 3.1\frac{\psi_2(1)}{\psi_3(1)} = \lambda_1 - 4
\]
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 \(\psi_2(1)/\psi_3(1)\)).

Mathematica gives (see WorkBook7, Section 2, Cell 2)
\[
\psi_1(1) = -0.542\psi_3(1) \quad (62)
\]
\[
\psi_2(1) = -0.914\psi_3(1) \quad (63)
\]
so the eigenvector \(\psi(1)\) is
\[
\psi(1) = \{-0.542\psi_3(1), -0.914\psi_3(1), \psi_3(1)\}
\]
\[
= \psi_3(1)\{-0.542, -0.914, 1\} \quad (64)
\]
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
\[
1 = \langle \psi(1) | \psi(1) \rangle = 2.13 |\psi_3(1)|^2 \quad (65)
\]
I solved Eq. 65 for $\psi_3(1)$ and I got 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:

$$
\psi_3(1) = 0.685
$$

(66)

and with this, I have

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

(67)

If you are the kind of person who worries about the mysteries of mathematics, you might wonder what happens if you use Eqs. 60 and 61, or perhaps Eqs. 59 and 61. I did this (see Workbook7, Section 2, Cell 2) and no matter which pair of equations I pick, I get the same result for $\psi_1(1)$ and $\psi_2(1)$! This must be so, if the theory is consistent, but it seems rather miraculous that it happens. However, no miracle is involved. Let us assume that we did not calculate the eigenvalue $\lambda$ and left it undetermined in Eqs. 59–61. We can take Eqs. 59 and 60 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 $\lambda$. Next we take Eqs. 59 and 61 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 $\lambda$. I calculate $\lambda$ from it and then I solve Eqs. 60 and 61, with this value of $\lambda$, for $\psi_1(1)/\psi_3(1)$ and $\psi_2(1)/\psi_3(1)$. The result I get must be equal to the one obtained previously. This demand for consistency of the system of homogeneous equations, Eqs. 59–61, is fulfilled only for special values of $\lambda$. These are the eigenvalues.

**Exercise 3** Implement the recipe outlined above for the homogeneous system

\[
(3.2 - \lambda)\psi_1 + 2.4\psi_2 = 0 \tag{68}
\]

\[
2.4\psi_1 + (3.6 - \lambda)\psi_2 = 0 \tag{69}
\]

Solve each equation for $\psi_1/\psi_2$ and demand that the ratio $(\psi_1/\psi_2)$ obtained from Eq. 68 be equal to the ratio obtained from Eq. 69. This gives you an equation for $\lambda$; 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 7.2. Degenerate eigenvalues

§ 10 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 could not find 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. Degenerate eigenvalues are fairly frequent and here I use two examples to suggest why.

§ 11 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 Hamiltonian is

$$\hat{H} = \hat{K}_x + \hat{K}_y + \hat{K}_z$$  \hspace{1cm} (70)

where

$$\hat{K}_x = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$$  \hspace{1cm} (71)

is the operator representing the kinetic energy of the motion in direction $x$; $\hat{K}_y$ and $\hat{K}_z$ are defined similarly. In classical mechanics, the total kinetic energy is

$$\frac{\mathbf{p}^2}{2m} = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}$$  \hspace{1cm} (72)

where $\mathbf{p}$ is the momentum vector and $p_x$ is the momentum in the direction $x$ (and analogously for $p_y$, $p_z$). The operator $\hat{K}_x$ is the quantum analog of $p_x^2/2m$.

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

$$E_{n,j,k} = \frac{\hbar^2 \pi^2}{2m} \left( \frac{n}{L_x} \right)^2 + \frac{\hbar^2 \pi^2}{2m} \left( \frac{j}{L_y} \right)^2 + \frac{\hbar^2 \pi^2}{2m} \left( \frac{k}{L_z} \right)^2$$  \hspace{1cm} (73)
where \( n, j, k \) can take any of the values

\[
n, j, k = 1, 2, 3, \ldots
\]  

(74)

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 \( n = 2, j = 1, k = 1 \) and the energy

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

(75)

**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 the cubic box, when

\[
L_x = L_y = L_z = L
\]  

(76)

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

\[
\frac{\hbar^2 \pi^2}{2mL^2} \left( 2^2 + 1 + 1 \right)
\]  

(77)

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)

$$\langle x, y, z \mid n, j, k \rangle \equiv \psi_{n,j,k}(x, y, z)$$

$$= \sqrt{\frac{2}{L_x}} \sin \left( \frac{n\pi x}{L_x} \right) \sqrt{\frac{2}{L_y}} \sin \left( \frac{j\pi y}{L_y} \right) \sqrt{\frac{2}{L_z}} \sin \left( \frac{k\pi z}{L_z} \right)$$

(78)

For the example of the cube,

$$\langle x, y, z \mid 2, 1, 1 \rangle = \sqrt{\frac{8}{L^3}} \sin \left( \frac{2\pi x}{L} \right) \sin \left( \frac{\pi y}{L} \right) \sin \left( \frac{\pi z}{L} \right)$$

(79)

and

$$\langle x, y, z \mid 1, 2, 1 \rangle = \sqrt{\frac{8}{L^3}} \sin \left( \frac{\pi x}{L} \right) \sin \left( \frac{2\pi y}{L} \right) \sin \left( \frac{\pi z}{L} \right)$$

(80)

A similar equation holds for $\langle x, y, z \mid 1, 1, 2 \rangle$. 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:

$$\hat{K}_x \langle x, y, z \mid 2, 1, 1 \rangle = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi_{2,1,1}(x, y, z)$$

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

(81)

I obtained the second line from the first by taking the second derivative of $\psi_{2,1,1}(x, y, z)$ given by Eq. 79. 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 $\frac{\hbar^2}{2m} \left( \frac{\pi}{L} \right)^2$. 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⟩, |1, 2, 1⟩, and |1, 1, 2⟩ 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 ($K_x$, $K_y$, $K_z$).

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 $L_z \neq L_y \neq L_z$?

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

$$E_n = -\frac{1}{n^2} \frac{\mu e^4}{2(4\pi \varepsilon_0)^2 \hbar^2}$$

and the states $|n, \ell, m\rangle$ are labeled by three quantum numbers: $n$ controls the energy, $\ell$ 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
\[ \hat{H}|n,\ell,m\rangle = E_n|n,\ell,m\rangle, \] (83) 
(2) the angular momentum squared $\hat{L}^2$, for which
\[ \hat{L}^2|n,\ell,m\rangle = \hbar^2(\ell+1)|n,\ell,m\rangle, \] (84) 
and (3) the projection $\hat{L}_z$ of the angular momentum on the OZ axis, for which
\[ \hat{L}_z|n,\ell,m\rangle = \hbar m|n,\ell,m\rangle, \] (85) 
The states $|2,0,0\rangle$, $|2,1,-1\rangle$, $|2,1,0\rangle$, and $|2,1,1\rangle$ (remember that $\ell$ takes only the values $0,1,\ldots,n-1$ and $m$ takes only the values $-\ell,-\ell+1,\ldots,\ell$) all have the energy
\[ E_2 = -\frac{1}{4} \frac{\mu e^4}{(2\pi\varepsilon_0)^2\hbar^2}. \] (86) 

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.

§ 13 Degenerate eigenvalues: an example. I hope you have accepted that degeneracy is not a freakish accident and you need to understand the mathematical properties of degenerate eigenstates. These are a bit more complicated than those of the non-degenerate states.

Let us start with an example. Mathematica tells me (see WorkBook7, Section 3, Cell 1) that the matrix
\[ M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \] (87)
has the eigenvalues
\[ \lambda_1 = 3 \]  
\[ \lambda_2 = 0 \]  
\[ \lambda_3 = 0 \]

with corresponding eigenvectors
\[ \mathbf{v}(1) = \{1, 1, 1\} \]  
\[ \mathbf{v}(2) = \{-1, 0, 1\} \]  
\[ \mathbf{v}(3) = \{-1, 1, 0\} \]

I calculated \( \langle \mathbf{v}(i) | \mathbf{v}(j) \rangle \) and found
\[ \langle \mathbf{v}(1) | \mathbf{v}(1) \rangle = 3, \quad \langle \mathbf{v}(2) | \mathbf{v}(2) \rangle = 2, \quad \langle \mathbf{v}(3) | \mathbf{v}(3) \rangle = 2 \]

The eigenvectors are \textit{not normalized}. I also calculated that
\[ \langle \mathbf{v}(1) | \mathbf{v}(2) \rangle = 0, \quad \langle \mathbf{v}(1) | \mathbf{v}(3) \rangle = 0, \quad \langle \mathbf{v}(2) | \mathbf{v}(3) \rangle = 1 \]

This means that the degenerate eigenvectors \( \mathbf{v}(2) \) and \( \mathbf{v}(3) \) are \textit{not orthogonal}. There are an infinite number of eigenvectors, and \texttt{Mathematica} has picked three of them by mysterious rules. They are not the best choice for physics, however, since we need them to be orthonormal.

We have seen earlier a theorem stating that the eigenvectors of a Hermitian matrix corresponding to \textit{different} eigenvalues are orthogonal. The matrix \( M \) is Hermitian, and \( \mathbf{v}(1) \) (corresponding to the eigenvalue 3) is orthogonal to \( \mathbf{v}(2) \) and \( \mathbf{v}(3) \) (corresponding to the eigenvalue 0). This is in agreement with the theorem. Since the eigenvalues corresponding to \( \mathbf{v}(2) \) and \( \mathbf{v}(3) \) are \textit{equal to each other}, \( \mathbf{v}(2) \) and \( \mathbf{v}(3) \) are \textit{under no obligation} to be orthogonal and we found that they are not. There is, however, a troublesome problem. If the Hermitian matrix represents an observable, we have assumed that its eigenstates form a complete, orthonormal system. They don’t seem to in this case but, as you’ll see soon, we can turn these eigenvectors into an orthonormal set of eigenvectors by using Gram-Schmidt orthogonalization.

\textbf{§ 14 Completeness.} Does this system provide a complete basis set in the space of three-dimensional vectors? I am asking whether it is possible to write any vector \( \mathbf{x} \) in our three-dimensional space in the form
\[ \mathbf{x} = \alpha_1 \mathbf{v}(1) + \alpha_2 \mathbf{v}(2) + \alpha_3 \mathbf{v}(3) \]
where $\alpha_1$, $\alpha_2$, $\alpha_3$ are suitably chosen numbers.

The answer will be yes if the three vectors are linearly independent, which means that there are no numbers $\beta_1$, $\beta_2$, $\beta_3$ such that

$$\beta_1 v(1) + \beta_2 v(2) + \beta_3 v(3) = 0 \quad (97)$$

(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. 97 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

$$B = \begin{pmatrix} v_1(1) & v_2(1) & v_3(1) \\ v_1(2) & v_2(2) & v_3(2) \\ v_1(3) & v_2(3) & v_3(1) \end{pmatrix} \quad (98)$$

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

$$\det B \neq 0 \quad (99)$$

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

$$\det B = -3 \quad (100)$$

We conclude that the three eigenvectors are linearly independent.

**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?

Because these three vectors are linearly independent, any three-dimensional vector $x$ can be written as

$$x = \mu_1 v(1) + \mu_2 v(2) + \mu_3 v(3) \quad (101)$$

where $\mu_1$, $\mu_2$, and $\mu_3$ are numbers. So $v(1)$, $v(2)$, $v(3)$ do form a complete basis set. But the three vectors making up the basis are not orthonormal.
Orthogonalization. We can use the Gram-Schmidt orthogonalization procedure (see Chapter 3, §17) to convert the degenerate eigenvectors \( \mathbf{v}(2) \) and \( \mathbf{v}(3) \) into two orthonormal vectors \( \mathbf{x}(2) \) and \( \mathbf{x}(3) \). These are

\[
\mathbf{x}(2) = \mathbf{v}(2) \quad (102)
\]

\[
\mathbf{x}(3) = \mathbf{v}(3) - \frac{\mathbf{x}(2)}{\langle \mathbf{x}(2) | \mathbf{x}(2) \rangle} \langle \mathbf{x}(2) | \mathbf{v}(3) \rangle \quad (103)
\]

The calculation was performed in Section 3, Cell 3 of WorkBook7 and the result is

\[
\mathbf{x}(2) = \mathbf{v}(2) = \{-1, 0, 1\} \quad (104)
\]

\[
\mathbf{x}(3) = \left\{-\frac{1}{2}, 1, -\frac{1}{2}\right\} \quad (105)
\]

It is easy to verify that \( M\mathbf{x}(2) = 0 = 0 \times \mathbf{x}(2) \) and \( M\mathbf{x}(3) = 0 = 0 \times \mathbf{x}(3) \), which means that \( \mathbf{x}(2) \) and \( \mathbf{x}(3) \) are eigenstates of \( M \) corresponding to the eigenvalue equal to zero. Also, we can verify that \( \langle \mathbf{v}(1) | \mathbf{x}(2) \rangle = \langle \mathbf{v}(1) | \mathbf{x}(3) \rangle = \langle \mathbf{x}(2) | \mathbf{x}(3) \rangle = 0 \); the vectors \( \mathbf{x}(i) \) are orthogonal, as expected.

Next we normalize \( \mathbf{x}(1), \mathbf{x}(2), \) and \( \mathbf{x}(3) \), by using

\[
s(1) = \frac{\mathbf{x}(1)}{\sqrt{\langle \mathbf{x}(1) | \mathbf{x}(1) \rangle}} \quad (106)
\]

\[
s(2) = \frac{\mathbf{x}(2)}{\sqrt{\langle \mathbf{x}(2) | \mathbf{x}(2) \rangle}} \quad (107)
\]

\[
s(3) = \frac{\mathbf{x}(3)}{\sqrt{\langle \mathbf{x}(3) | \mathbf{x}(3) \rangle}} \quad (108)
\]

The orthonormal vectors are therefore (see WorkBook7)

\[
s(1) = \left\{ \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right\} \quad (109)
\]

\[
s(2) = \left\{ -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right\} \quad (110)
\]

\[
s(3) = \left\{ -\frac{1}{\sqrt{6}} \sqrt{\frac{2}{3}}, -\frac{1}{\sqrt{6}} \right\} \quad (111)
\]
Exercise 8 Verify that (1) \( Ms(i) = \lambda s(i) \), \( i = 1, 2, 3 \), 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).

§ 16 Find the eigenvalues of the matrix \( M \) (Eq. 87) “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

\[
Mx = \lambda x
\]

For the matrix given by Eq. 87, this is the same as the system of equations

\[
(1 - \lambda)x_1 + x_2 + x_3 = 0 \quad (113)
\]
\[
x_1 + (1 - \lambda)x_2 + x_3 = 0 \quad (114)
\]
\[
x_1 + x_2 + (1 - \lambda)x_3 = 0 \quad (115)
\]

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

The characteristic polynomial is

\[
\det[M - \lambda I] = \det \begin{pmatrix} 1 - \lambda & 1 & 1 \\ 1 & 1 - \lambda & 1 \\ 1 & 1 & 1 - \lambda \end{pmatrix} = 3\lambda^2 - \lambda^3 \quad (116)
\]
The eigenvalues are the roots of this polynomial, i.e. the solutions of the equation
\[ 3\lambda^2 - \lambda^3 = 0 \] (117)

They are (see **WorkBook7.nb**, Section 3, Cell 4)
\[\begin{align*}
\lambda_1 &= 3 \\
\lambda_2 &= 0 \\
\lambda_3 &= 0
\end{align*}\] (118)(119)(120)

We can calculate the eigenvector \(x(1)\) corresponding to \(\lambda_1 = 3\) as we did before (this eigenvector is not degenerate). The equation \((M - \lambda_1 I)x(1) = 0\) is shorthand for the system of linear equations:
\[\begin{align*}
(1 - 3)x_1(1) + x_2(1) + x_3(1) &= 0 \\
x_1(1) + (1 - 3)x_2(1) + x_3(1) &= 0 \\
x_1(1) + x_2(1) + (1 - 3)x_3(1) &= 0
\end{align*}\] (121)(122)(123)

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. 116 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. 121 and 122 and obtained
\[\begin{align*}
x_2(1) &= x_1(1) \\
x_3(1) &= x_1(1)
\end{align*}\] (124)(125)

The eigenvector \(x(1)\) corresponding to the eigenvalue \(\lambda_1 = 3\) is therefore
\[x(1) = \{x_1(1), x_1(1), x_1(1)\} = x_1(1)\{1, 1, 1\}\] (126)

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
\[s(1) = \frac{x(1)}{\sqrt{(x(1)|x(1))}} = \left\{ \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right\}\] (127)

Notice that \(x_1(1)\) disappears from this expression.
We have

\[(M - 0I)x(2) = 0\]  \hspace{1cm} (128)

and

\[(M - 0I)x(3) = 0\]  \hspace{1cm} (129)

because the eigenvalues \(\lambda_2\) and \(\lambda_3\) are equal to zero. However, we must have \(x(2) \neq x(3)\) because they are different eigenstates. This can happen only if all three equations in the system \((M - 0I)x = 0\) are identical. If they were not, we could proceed as we did when we calculated \(x(1)\). Take two of the equations in the system \((M - 0I)x = 0\) and solve for \(x(2)\) to get a solution of the form

\[x(2) = x_3(2)\{a, b, 1\}\]

where \(a\) and \(b\) are numbers whose values are irrelevant for this discussion. Normalizing this vector fixes \(x_3(2)\). But the system for \(x(3)\) is the same as that for \(x(2)\). Therefore when we solve it we obtain \(x(3) = x(2)\). This is not acceptable and therefore all three equations in the system must be the same. Indeed when we insert \(\lambda_2 = 0\) in Eqs. 113–115 we obtain

\[x_1(2) + x_2(2) + x_3(2) = 0\]  \hspace{1cm} (130)
\[x_1(2) + x_2(2) + x_3(2) = 0\]  \hspace{1cm} (131)
\[x_1(2) + x_2(2) + x_3(2) = 0\]  \hspace{1cm} (132)

The equations for \(x_1(3), x_2(3),\) and \(x_3(3)\) are identical to Eqs. 130–132 because \(\lambda_2 = \lambda_3\). Inserting \(\lambda_3 = 0\) in Eqs. 113–115 yields three identical equations (i.e. Eqs. 130–132).

All I can conclude from Eqs. 130–132 is that

\[x_1(2) = -x_2(2) - x_3(2)\]  \hspace{1cm} (133)

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

\[x(\alpha) = \{-x_2(\alpha) - x_3(\alpha), x_2(\alpha), x_3(\alpha)\}, \ \alpha = 2, 3\]  \hspace{1cm} (134)

Mathematics does not further determine \(x_2\) and \(x_3\); any vector of this form satisfies the equation \(Mx = 0x\), as you can easily verify. In addition, any vector of the form in Eq. 134 is orthogonal to \(s(1)\) (see WorkBook7) because \(s(1)\) corresponds to a different eigenvalue of \(M\).
Exercise 9 For \( x = \{ -a - b, a, b \} \), test that \( Mx = 0 \) and \( x \cdot s(1) = 0 \).

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 in Eq. 134)

\[
    x(2) = \{-3, 1, 2\}
\]

(135)

Then I pick a second vector by choosing arbitrarily\(^1\) \( x_2(3) = 0 \) and \( x_3(3) = 1 \) and obtain

\[
    x(3) = \{-1, 0, 1\}
\]

(136)

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

\[
    v(2) = x(2) = \{-3, 1, 2\}
\]

(137)

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

(138)

Exercise 10 Show that \( \langle s(1) | v(2) \rangle = \langle s(1) | v(3) \rangle = \langle v(2) | v(3) \rangle = 0 \) and \( Mv(i) = 0v(i) \) for \( i = 2, 3 \).

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

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

(139)

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

(140)

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.

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

\(^1\)or almost so: make sure not to choose a multiple of the first vector
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 $9 \times 9$ matrix and an eigenvalue is six-fold degenerate, you lose seven equations; and so forth.

Appendix 7.1. A few properties of determinants

1. The determinant of a Hermitian or a unitary matrix is equal to the product of its eigenvalues. Therefore the determinant of a Hermitian matrix is a real number (even when the matrix has complex elements) and the absolute value of the determinant of a unitary matrix is equal to 1.

2. We have
   \[
   \det(A^{-1}) = \frac{1}{\det A}
   \]  
   (141)
   where by writing $A^{-1}$ we are asserting that it exists.

3. The determinant of a product is the product of the determinants:
   \[
   \det(AB) = (\det A)(\det B)
   \]  
   (142)

4. If a row of a matrix is a linear combination of other rows, then the determinant of the matrix is zero.

5. If a column of a matrix is a linear combination of other columns, then the determinant of the matrix is zero.

Exercise 11 Prove Eq. 141 when $A$ is the matrix representation of $\hat{A} = \sum_{\alpha} |\alpha\rangle \langle \alpha|.$

Exercise 12 Show that $\det(A^{-1}BA) = \det B.$
Appendix 7.2. Matrix diagonalization

The following theorem is often used to simplify proofs and calculations. *Theorem.* For any Hermitian matrix $\hat{H}$ there is a unitary matrix $U$ such that

$$ U^{-1}\hat{H}U \text{ is a diagonal matrix } \Lambda $$

and the eigenvalues of $H$ are the diagonal elements of $\Lambda$.

Whenever we talk about equations involving matrices, their elements are calculated with the same basis set $\{|e(i)\rangle\}_{i=1}^{N}$. Since $\hat{H}$ is Hermitian, it has a set of eigenkets $\{|x(i)\rangle\}_{i=1}^{N}$ that satisfy the eigenvalue equation

$$ \hat{H}|x(i)\rangle = \lambda_i |x(i)\rangle, \quad i = 1, \ldots, N \quad (144) $$

Define the operator $\hat{U}$ through

$$ |x(i)\rangle \equiv \hat{U}|e(i)\rangle \quad (145) $$

Because $\{|e(i)\rangle\}_{i=1}^{N}$ and $\{|x(i)\rangle\}_{i=1}^{N}$ are both orthonormal, complete basis sets, $\hat{U}$ is unitary (see Property 14 in Chapter 4). Replace $|x(i)\rangle$ in the eigenvalue equation (Eq. 144) with $\hat{U}|e(i)\rangle$ and then apply $\hat{U}^{-1}$ to both sides of the resulting equation. The result is

$$ \hat{U}^{-1}\hat{H}\hat{U}|e(i)\rangle = \lambda_i |e(i)\rangle, \quad i = 1, \ldots, N \quad (146) $$

Act on this with $\langle e(j)|$ and use $\langle e(j) | e(i) \rangle = \delta_{ji}$ to obtain

$$ \langle e(j)| \hat{U}^{-1}\hat{H}\hat{U}|e(i)\rangle = \lambda_i \delta_{ji} \equiv \Lambda_{ji}, \quad i, j = 1, \ldots, N \quad (147) $$

The expression $\langle e(j)| \hat{U}^{-1}\hat{H}\hat{U}|e(i)\rangle$ is the matrix element of the operator $\hat{U}^{-1}\hat{H}\hat{U}$ in the $\{|e(j)\rangle\}$ representation. The matrix $\Lambda$ (having elements $\Lambda_{ji} = \lambda_i \delta_{ji}$) is diagonal:

$$ \Lambda = \begin{pmatrix}
\lambda_1 & 0 & 0 & \cdots & 0 \\
0 & \lambda_2 & 0 & \cdots & 0 \\
0 & 0 & \lambda_3 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \lambda_N
\end{pmatrix} \quad (148) $$
The left-hand side of Eq. 147 can be written as (use $\sum_P |e(p)\rangle \langle e(p)| = \hat{I}$)

$$\sum_{p=1}^{N} \sum_{k=1}^{N} \langle e(j) | \hat{U}^{-1} | e(p)\rangle \langle e(p) | \hat{H} | e(k)\rangle \langle e(k) | \hat{U} | e(i)\rangle = \sum_{p=1}^{N} \sum_{k=1}^{N} (U^{-1})_{jp} H_{pk} U_{ki} = \left(U^{-1} H U\right)_{ji}$$

(149)

where we have denoted the matrix elements by

$$(U^{-1})_{jp} \equiv \langle e(j) | \hat{U}^{-1} | e(p)\rangle$$

$$H_{pk} \equiv \langle e(p) | \hat{H} | e(k)\rangle$$

$$U_{ki} \equiv \langle e(k) | \hat{U} | e(i)\rangle$$

and used the observation that the second double sum in Eq. 149 is the rule for matrix multiplication. This allows us to write the last term (i.e. $U^{-1} H U$) as the product of the matrices $U^{-1}$, $H$, and $U$. Combining Eq. 147 with Eq. 149 proves the theorem.

We can go a step further and construct the matrix $U$ from the eigenvectors of $\hat{H}$. We have defined $\hat{U}$ through Eq. 145,

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

Acting with $\langle e(j)|$ on this gives

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

(150)

Using the completeness relation for the basis set $\{|e(i)\rangle\}_{i=1}^{N}$, we have

$$|x(i)\rangle = \sum_{j=1}^{N} \langle e(j) | e(i)\rangle |e(j)\rangle$$

(151)

This means that $\langle e(j) | x(i)\rangle$ are the components of the eigenket $|x(i)\rangle$ in the basis $\{|e(i)\rangle\}_{i=1}^{N}$. They are the numbers we obtain when we find the eigenvectors of the matrix $H$; the eigenvectors are $x(i) = \{\langle e(1) | x(i)\rangle, \langle e(2) | x(i)\rangle, \ldots, \langle e(N) | x(i)\rangle\}$. The matrix $U$ is (see Eq. 150)

$$U = \begin{pmatrix}
\langle e(1) | x(1)\rangle & \langle e(1) | x(2)\rangle & \cdots & \langle e(1) | x(N)\rangle \\
\langle e(2) | x(1)\rangle & \langle e(2) | x(2)\rangle & \cdots & \langle e(2) | x(N)\rangle \\
\langle e(3) | x(1)\rangle & \langle e(3) | x(2)\rangle & \cdots & \langle e(3) | x(N)\rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle e(N) | x(1)\rangle & \langle e(N) | x(2)\rangle & \cdots & \langle e(N) | x(N)\rangle
\end{pmatrix}$$

(152)
Remember that if $U_{ij}$ is a matrix element, the first index (here, $i$) labels the rows and the second index labels the columns. A simple way to construct $U$ according to Eq. 152 is to take the first eigenvector of $\hat{H}$ and use it as the first column in $U$, then use the second eigenvector as the second column of $U$, etc.

To construct $U$, we need to solve the eigenvalue problem for $H$. Because of this, constructing $U$ is not a shortcut for finding the eigenvectors (from Eq. 145) or the eigenvalues (from Eq. 143). The theorem is however very useful for simplifying some equations and as an intermediate step in some mathematical proofs.

In Section 4, Cells 6–7, of the file \texttt{linear algebra for quantum mechanics.nb}, I give an example of the construction of a unitary matrix $U$ that diagonalizes a Hermitian matrix $M$. In the following displays, the numbers have been rounded to two significant digits.

$$M = \begin{pmatrix}
-1.1 & 2.3 + 0.022i & -0.67 + 3.0i & -3.5 + 5.0i \\
2.3 - 0.022i & -2.0 & 4.2 + 3.9i & -1.1 + 3.1i \\
-0.67 - 3.0i & 4.2 - 3.9i & 0.27 & -2.5 + 4.6i \\
-3.5 - 5.0i & -1.1 - 3.1i & -2.5 - 4.6i & 1.2
\end{pmatrix} \quad (153)$$

The eigenvectors are (see Cell 7 of \texttt{linear algebra for quantum mechanics.nb})

$$x(1) = \{-0.33 + 0.26i, -0.26 + 0.34i, -0.024 + 0.53i, 0.60\} \quad (154)$$
$$x(2) = \{-0.25 + 0.42i, 0.40 + 0.18i, -0.50 + 0.26i, -0.50\} \quad (155)$$
$$x(3) = \{-0.026 - 0.50i, 0.74 + 0.024i, -0.16 + 0.11i, 0.41\} \quad (156)$$
$$x(4) = \{-0.18 - 0.54i, -0.12 + 0.26i, 0.36 + 0.49i, -0.47\} \quad (157)$$

Construct $U$ by using the eigenvectors as columns:

$$U = \begin{pmatrix}
-0.33 + 0.26i & -0.25 + 0.42i & -0.026 - 0.50i & -0.18 - 0.54i \\
-0.26 + 0.34i & 0.40 + 0.18i & 0.74 + 0.024i & -0.12 + 0.26i \\
-0.024 + 0.53i & -0.50 + 0.26i & -0.16 + 0.11i & 0.36 + 0.49i \\
0.60 & -0.50 & 0.41 & -0.47
\end{pmatrix} \quad (158)$$

We calculate that (see \texttt{linear algebra for quantum mechanics.nb})

$$U^{-1}MU = \begin{pmatrix}
11.63 & 0 & 0 & 0 \\
0 & -9.87 & 0 & 0 \\
0 & 0 & -4.16 & 0 \\
0 & 0 & 0 & 0.77
\end{pmatrix} \quad (159)$$
The numbers on the diagonal are the eigenvalues of $M$.

**Exercise 13** Show that if $M$ is Hermitian and $A$ has an inverse, then $A^{-1}MA$ and $AMA^{-1}$ have the same eigenvalues as $M$. Try to find at least two methods of proof.