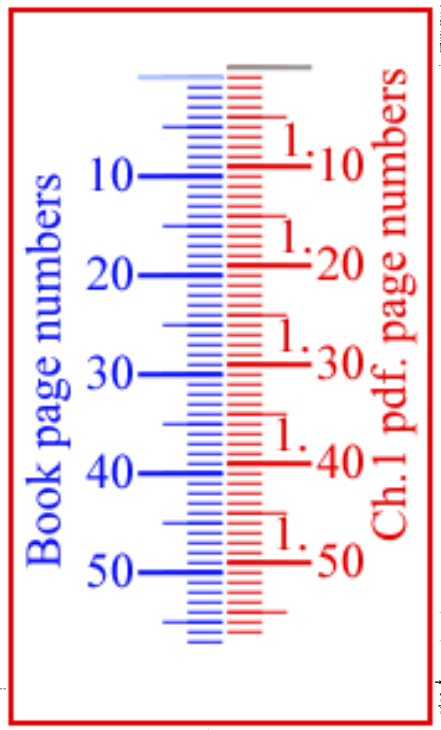
Principles of Symmetry, Dynamics, and Spectroscopy
- W. G. Harter - Wiley (1993)

Chapter



CHAPTER 1

A REVIEW OF MATRIX ALGEBRA AND QUANTUM MECHANICS

The study of symmetry and its application to spectroscopy involves the mathematics of operators, vectors, and matrices, and this chapter contains a review of matrix theory and some of its applications. The review includes a few concepts and procedures which may not be well known to physicists, but which are needed in the development of this book. Also, we use the review to establish much of our symbolism, conventions, and definitions. Finally, there are some previews of the contents of the rest of the book.

1.1 MATRICES USED IN QUANTUM MECHANICS

The matrix description of quantum mechanics was first developed by Heisenberg and has since become widely used. This description differs from some others mostly by the organization of its bookkeeping, wherein numbers of physical interest are stored in arrays called MATRICES. Such an array is shown in Eq. (1.1.1):

$$\mathscr{M} = \begin{pmatrix} \mathscr{M}_{11} & \mathscr{M}_{12} & \mathscr{M}_{13} & \cdots \\ \mathscr{M}_{21} & \mathscr{M}_{22} & \mathscr{M}_{23} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}. \tag{1.1.1}$$

We will use script notation $\mathscr{M}, \mathscr{N}, \ldots$ for a whole matrix, and let the same letter with subscripts $\mathscr{M}_{ij}$ designate the (complex) number which is the COMPONENT of the matrix, located where the ith row meets the jth column.

Many, but not all, of the matrices which we shall use will be finite-square or $(n \times n)$ matrices, that is, matrices with n rows and n columns, where $n = 1, 2, 3, \ldots$ is finite. [A (1×1) matrix is just a number.] This is because many physical situations can be described by some relation between a set of n quantum states and some other set of n different states. Consider now some examples of such matrices.

A. Transformation Matrix

Quite often in quantum mechanics we can describe a particle or a system of particles by stating whether they are in one or another of some set of n quantum states $1, 2, 3, \ldots, n$. The most famous example of this might be the states (1): spin "up" and (2): spin "down" for an electron.

Now someone else can come along and describe the same system with a different set of states $1', 2', 3', \ldots,$ and n', which are just as complete in their description as the first set. For example, 1': spin "north" and 2': spin "south" are acceptable states for the electron, too, as we will see shortly.

However, you must know the relation between any two equivalent descriptions. This relation is given entirely by a TRANSFORMATION MATRIX $\mathscr{T}(b' \leftarrow b)$, the meaning of which we shall review:

$$\mathscr{T}(b' \leftarrow b) = \begin{pmatrix} \langle 1'|1\rangle & \langle 1'|2\rangle & \cdots & \langle 1'|n\rangle \\ \langle 2'|1\rangle & \langle 2'|2\rangle & \cdots & \langle 2'|n\rangle \\ \vdots & \vdots & & \\ \langle n'|1\rangle & \langle n'|2\rangle & \cdots & \langle n'|n\rangle \end{pmatrix}. \tag{1.1.2}$$

For the time being we shall exhibit the $\mathscr{T}$ matrix for going between "up-down" (1-2) and "north-south" (1'-2') directions shown in Figure 1.1.1 for electronic spin. (The figure shows up-down tilted to avoid "geo-chauvinism" or undue favoring of a particular inhabited latitude.) The $\mathscr{T}$ matrix that follows and generalizations of it are derived in Chapter 5, and in fact most of symmetry theory is devoted to finding some sort of $\mathscr{T}$ matrix:

$$\mathscr{T}(b' \leftarrow b) = \begin{pmatrix} \langle 1'|1\rangle & \langle 1'|2\rangle \\ \langle 2'|1\rangle & \langle 2'|2\rangle \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}. \tag{1.1.3}$$

It is important to understand the meaning of a $\mathscr{T}$ matrix. It is conventional to say that component $\langle i'|j\rangle$ gives the "PROBABILITY AMPLITUDE for a system in state j to be found in state i'...". Now one has a perfect right to ask what that means, but many students find it difficult to get a satisfying answer. However, it is not so difficult to learn that $\langle i'|j\rangle$ *squared*,

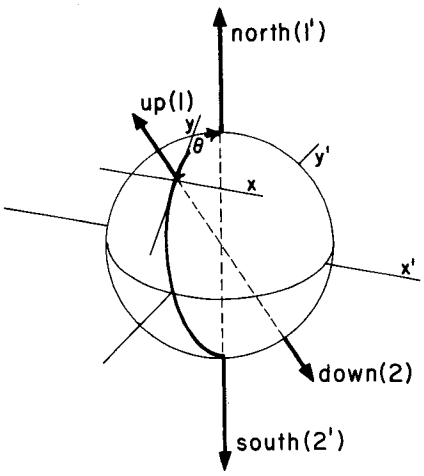


Figure 1.1.1 Alternative axes for spin state definition. The up-down axis (z) is rotated by angle θ from the north-south axis (z').

i.e., $|\langle i'|j\rangle|^2$, is the fraction of, or probability for, systems in state j to end up in state i' when forced to make the choice. The difficulty is that it is apparently impossible to predict the outcome of an individual event. To know this with certainty we must really wait for the individual particle or system to make its choice. The prediction $|\langle i'|j\rangle|^2$ is only the probability for the individual outcome [(i') from (j)] or the average percentage for choices in a large number of individual experiments.

A device that forces a particle to make a choice is called an *analyzer*. An example of a spin-state analyzer is the magnetic Stern-Gerlach device which Feynman has described in his *Lectures on Physics* (Vol. III, p. 5-1). The analyzer accepts a beam of spinning particles and displaces each one up or down according to the projection of the particles' spin on the vertical axis of the analyzer. For electrons only two possible projections are possible. One possible projection is along the axis which is up [state (1)] for an up-down analyzer, or north [state (1')] for a north-south pointing analyzer, and so on. The other possibility is opposite to the axis, i.e., down [state (2)], or south [state (2')], etc. Given an incoming beam of particles in state (j), one finds *on the average* that a fraction $|\langle i'|j\rangle|^2 = |\mathscr{T}_{i'j}|^2$ of them will come out of the (i') exit of the (1')-(2') or north-south analyzer.

For example, the fraction of spin-up (1) electrons that would actually choose to point south (2') in a north-south analyzer (see Figure 1.1.2) is $|\langle 2'|1\rangle|^2 = \sin^2(\theta/2)$ given $\langle 2'|1\rangle$ in Eq. (1.1.3). The remaining fraction $1 - \sin^2(\theta/2) = \cos^2(\theta/2)$ end up choosing north (1').

Understanding the amplitude $\langle i'|j\rangle$, apart from its square, is more difficult. Let us state some axioms which most people like to have the amplitudes obey, and discuss some thought experiments which may help us to make them more plausible. The first axiom has already been introduced.

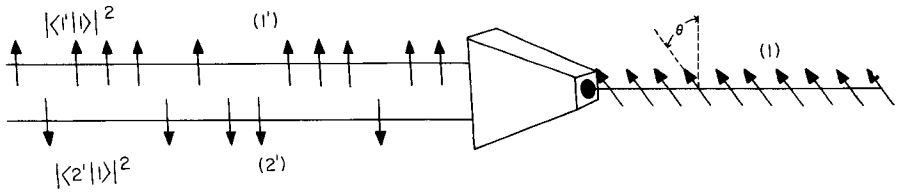


Figure 1.1.2 Spin analyzer experiment. An incident beam of particles in state (up) comes in on the right-hand side. The north-south analyzer forces each particle to choose between state 1′ (north) or state 2′ (south). On the average a fraction $|\langle 2'|1\rangle|^2 = \sin^2(\theta/2)$ choose the latter state 2′ and emerge from the bottom of the analyzer. The remaining fraction $|\langle 1'|1\rangle|^2 = \cos^2(\theta/2)$ come out at the top in the state 1′.

Axiom 1 $\langle i'|j\rangle^* \langle i'|j\rangle = |\langle i'|j\rangle|^2$ is the probability for occurrence in state i' of a system originally in state j which is forced to choose between $1', 2', \ldots, i', \ldots$.

One should note that in classical wave and polarization theory the absolute square ($|A|^2$) of an amplitude denotes *intensity*. In quantum theory it gives probability, which is analogous.

The second axiom deals with complex conjugation ($*$) and order reversal.

Axiom 2 $\langle i'|j\rangle^* = \langle j|i'\rangle$.

In Axiom 2 one supposes that a reverse process obtained by playing time backwards is accomplished mathematically by complex conjugation. For example, a plane-wave amplitude $\langle x, t|k, \omega\rangle = e^{i(kx-\omega t)}$ reversed is $\langle x, t|k, \omega\rangle^* = e^{i([-k]x-\omega[-t])}$. This axiom is a direct carryover from classical wave theory.

The next axiom tells how states within a given set are related.

Axiom 3 *The amplitudes connecting a set of states* $\{1, 2, \ldots, n\}$ *with itself* are 1 or 0; that is, $\langle i|j\rangle = \delta_{ij} \equiv \begin{Bmatrix} 1 & i=j \\ 0 & i\neq j \end{Bmatrix}$.

Axiom 3 is closely related to the following experimental result. If some electrons that originally chose the state of spin north (1′) are once again given the chance in another analyzer to choose north or south (Figure 1.1.3), then 100% will choose north. This means $|\langle 1'|1'\rangle|^2 = 1$ and $|\langle 2'|1'\rangle|^2 = 0$. One then assumes that the phases are such that $\langle 1'|1'\rangle = 1$. The same applies to the (2′) state.

Axiom 3 means that the transformation matrix of states $\{1, 2, \ldots, n\}$ to themselves, or any other set $\{1', 2', \ldots, n\}$ to itself must be the IDENTITY or

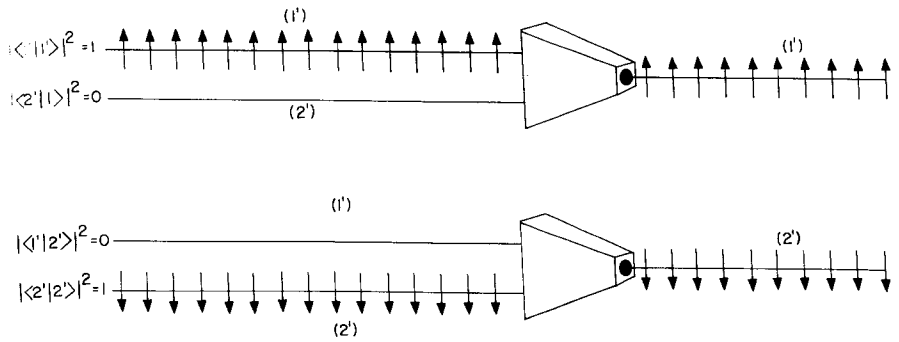


Figure 1.1.3 Analyzer experiment associated with Axiom 3. If the incident beam contains only particles in state 1′ (north) or separate (south) then they will be unchanged by a north-south analyzer.

UNIT MATRIX, *1*:

$$\begin{pmatrix} \langle 1|1\rangle & \langle 1|2\rangle & \cdots \\ \langle 2|1\rangle & \langle 2|2\rangle & \cdots \\ \vdots & \vdots & \end{pmatrix} = \begin{pmatrix} \langle 1'|1'\rangle & \langle 1'|2'\rangle & \cdots \\ \langle 2'|1'\rangle & \langle 2'|2'\rangle & \cdots \\ \vdots & \vdots & \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots \\ 0 & 1 & \cdots \\ \vdots & \vdots & \end{pmatrix} = 1. \tag{1.1.4}$$

The conditions $\langle i|j\rangle = \delta_{ij}$ are often called the ORTHONORMALITY conditions for these states, as we shall explain later.

The fourth and final axiom contains the crucial aspects of quantum wave mechanics.

Axiom 4 Given three sets of states $\{1, 2, \ldots, n\}$, $\{1', 2', \ldots, n'\}$, and $\{1'', 2'', \ldots, n''\}$ all relations of the following form must hold between their amplitudes:

$$\langle k''|i'\rangle = \sum_{j=1}^{n} \langle k''|j\rangle\langle j|i'\rangle.$$

Axiom 4 may be hard to swallow, but at least it is easy to see that it is consistent with Axioms 1, 2, and 3. A sum of probabilities should be unity if all states are accounted for, and according to 1 and 2 this sum is the following:

$$1 = \sum_{i'=1}^{n} |\langle i'|j\rangle|^2 = \sum_{i'=1}^{n} \langle j|i'\rangle\langle i'|j\rangle, \tag{1.1.5a}$$

and by Axiom 4 this equals

$$1 = \langle j|j\rangle, \tag{1.1.5b}$$

which checks with Axiom 3. However, one probably needs more than this to trust Axiom 4.

Two analyzer experiments meant to test Axiom 4 are depicted in Figures 1.1.4(a) and 1.1.4(b). The idea is that a beam of electrons all in a definite state (1′) [Figure 1.1.4(a)] will not be affected by a separation into some other states, say (1) and (2), if this is followed by a *coherent* recombination as indicated in the center of Figure 1.1.4(b). [A time-reversed analyzer indicated by a box whose direction is reversed in the figure is used to reform the particles into a single beam.] The two experiments shown give the same *average* number of electrons in the final states (1″), (2″) on the left, if the recombination is coherent. If the recombination is incoherent, then the average number of electrons in final state (k'') would approach the sum of squares ($\Sigma|\langle k''|j\rangle\langle j|1'\rangle|^2$) instead of the square of the sum ($|\Sigma\langle k''|j\rangle\langle j|1'\rangle|^2$). This would happen if some device in the (1-2) analyzer determined whether each electron went through in the (1) or in the (2) state. This is because the device would add a random phase to amplitude $\langle k''|j\rangle$, i.e., a different phase for each particle that went through. Then only the positive-definite terms in the square of the sum

$$\left|\sum\langle k''|j\rangle\langle j|1'\rangle\right|^2 = \sum \text{(positive definite)} + \sum \text{(phase sensitive)}$$

would survive after averaging over many particles, i.e., just the sum of squares or the sum of probabilities for each j:

$$\sum \text{(positive definite)} = \sum_j |\langle k''|j\rangle\langle j|1'\rangle|^2.$$

Meanwhile, the phase-sensitive "interference" terms

$$\sum \text{(phase sensitive)} = \sum_{j\neq k}\sum (\langle k''|j\rangle\langle j|1'\rangle)^*(\langle k''|k\rangle\langle k|1'\rangle)$$

would average to zero.

Axiom 4 can also be visualized in terms of basic wave mechanics. It is essentially a restatement of the principle of Huygens (circa 1660). Imagine a lightwave in state i' (Figure 1.1.5), going through a film to be (possibly) in some state k''. If each of the points j on the film absorb the light, but then rebroadcast it starting with the right amplitude and phase $\langle j|i'\rangle$ for that point, then Axiom 4 says state k'' will be achieved just as well as it would have if the film had not been there at all. This principle will be discussed in Chapter 8.

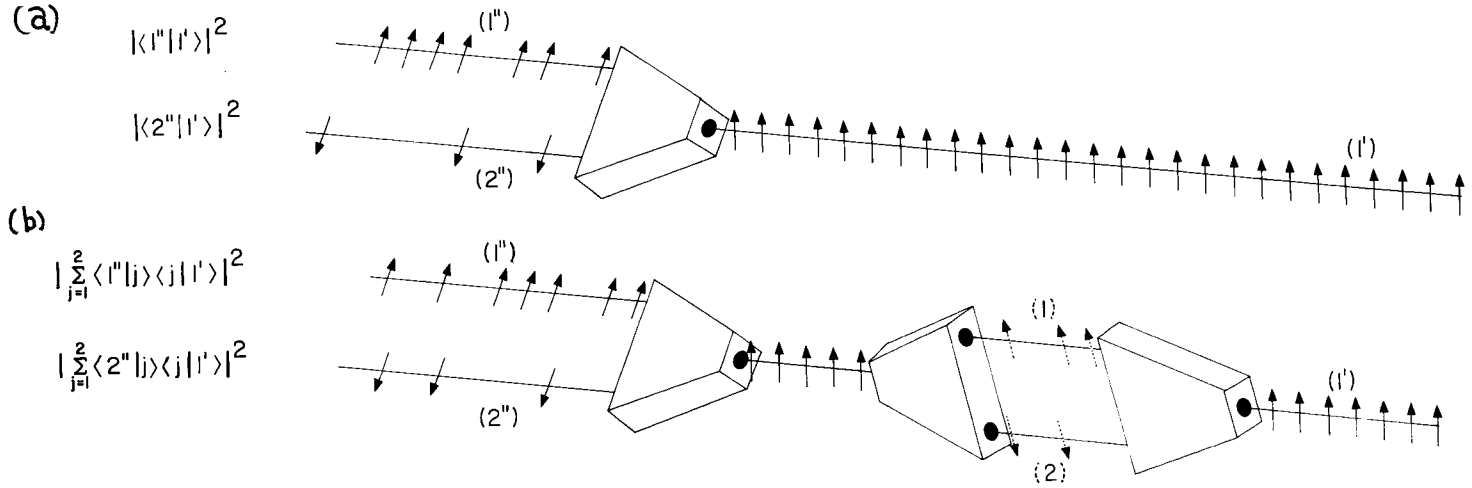


Figure 1.1.4 Analyzer experiments associated with Axiom 4. (a) Undisturbed beam in state 1′ (north) is analyzed by a 1″-2″ analyzer. (b) The second experiment is the same as the first except the 1′ beam is first analyzed by a 1-2 analyzer followed by a reversed (∗) analyzer which recombines the split beam. Axiom 4 says that such an arrangement will not effect the beam.

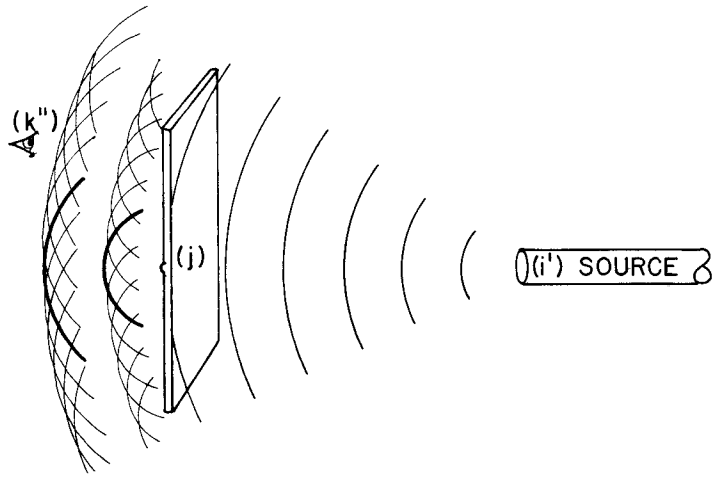


Figure 1.1.5 Relating Huygen's principle to Axiom 4.

Axiom 4 is the same as mathematical relations called the COMPLETENESS CONDITIONS for states $\{1, 2, \ldots, n\}$, and we shall soon discuss these from several points of view. Such relations are central to the development of symmetry analysis.

Furthermore, Axiom 4 is the real reason for ever studying matrix mathematics in the first place. Let us rewrite Axiom 4 using the script notation introduced in Eq. (1.1.2). Axiom 4 then corresponds to the standard definition of INNER MATRIX MULTIPLICATION (we discuss outer multiplication later):

$$\mathscr{T}_{k''i'}(b'' \leftarrow b') = \sum_{j=1}^{n} \mathscr{T}_{k''j}(b'' \leftarrow b)\mathscr{T}_{ji'}(b \leftarrow b'), \tag{1.1.6a}$$

$$\mathscr{T}(b'' \leftarrow b') = \mathscr{T}(b'' \leftarrow b)\mathscr{T}(b \leftarrow b'). \tag{1.1.6b}$$

Just writing two script letters next to each other, as in Eq. (1.1.6b), implies that the special sum in Eq. (1.1.6a) is to be performed.

Continuing the review of matrix mathematics, we define the TRANSPOSE CONJUGATE† of matrix $\mathscr{M}$ by

$$\mathscr{M}^\dagger = \begin{pmatrix} \mathscr{M}_{11} & \mathscr{M}_{12} & \cdots & \mathscr{M}_{1n} \\ \mathscr{M}_{21} & \mathscr{M}_{22} & \cdots & \mathscr{M}_{2n} \\ \vdots & \vdots & & \\ \mathscr{M}_{n1} & \mathscr{M}_{n2} & \cdots & \mathscr{M}_{nn} \end{pmatrix}^\dagger = \begin{pmatrix} \mathscr{M}_{11}^* & \mathscr{M}_{21}^* & \cdots & \mathscr{M}_{n1}^* \\ \mathscr{M}_{12}^* & \mathscr{M}_{22}^* & \cdots & \mathscr{M}_{n2}^* \\ \vdots & \vdots & & \\ \mathscr{M}_{1n}^* & \mathscr{M}_{2n}^* & \cdots & \mathscr{M}_{nn}^* \end{pmatrix} \tag{1.1.7}$$

Observe that the transformation matrix $\mathscr{T}$ satisfies Eq. (1.1.8), which is written in several different forms, starting with Axiom 4:

$$\sum_{i'=1}^{n} \langle k|i'\rangle\langle i'|j\rangle = \delta_{kj}, \qquad \sum_{i=1}^{n} \langle j'|i\rangle\langle i|k'\rangle = \delta_{k'j'} \tag{1.1.8a}$$

$$\sum_{i'=1}^{n} \mathscr{T}^*_{i'k}(b' \leftarrow b)\mathscr{T}_{i'j}(b' \leftarrow b) = \delta_{kj}, \qquad \sum_{i=1}^{n} \mathscr{T}_{j'i}(b' \leftarrow b)\mathscr{T}^*_{k'i}(b' \leftarrow b) = \delta_{k'j'}, \tag{1.1.8b}$$

$$\mathscr{T}^\dagger(b' \leftarrow b)\mathscr{T}(b' \leftarrow b) = 1 = \mathscr{T}(b' \leftarrow b)\mathscr{T}^\dagger(b' \leftarrow b). \tag{1.1.8c}$$

Any matrix satisfying an equation of the last form, Eq. (1.1.8c), is called a UNITARY MATRIX. In fact, when the product of any two finite matrices $\mathscr{A}$ and $\mathscr{B}$ is equal to 1, then each is said to be the INVERSE of the other, written $\mathscr{A} = \mathscr{B}^{-1}$ or $\mathscr{B} = \mathscr{A}^{-1}$ (Appendix B reviews the calculation of inverses.) The unitary matrix has the very convenient property that its inverse is just its transpose conjugate ($\mathscr{T}^{-1} = \mathscr{T}^\dagger$).

For any real matrix $\mathscr{R}$ ($\mathscr{R}$ has no complex components) the transpose conjugate $\mathscr{R}^\dagger$ simply equals the TRANSPOSE $\mathscr{R}^T$ defined by

$$\mathscr{R}^T = \begin{pmatrix} \mathscr{R}_{11} & \mathscr{R}_{12} & \cdots & \mathscr{R}_{1n} \\ \mathscr{R}_{21} & \mathscr{R}_{22} & \cdots & \mathscr{R}_{2n} \\ \mathscr{R}_{n1} & \mathscr{R}_{n2} & \cdots & \mathscr{R}_{nn} \end{pmatrix}^T \equiv \begin{pmatrix} \mathscr{R}_{11} & \mathscr{R}_{21} & \cdots & \mathscr{R}_{n1} \\ \mathscr{R}_{12} & \mathscr{R}_{22} & \cdots & \mathscr{R}_{n2} \\ \mathscr{R}_{1n} & \mathscr{R}_{2n} & \cdots & \mathscr{R}_{nn} \end{pmatrix}. \tag{1.1.9}$$

Any real transformation matrix will satisfy Eq. (1.1.10). A matrix satisfying this equation is said to be an ORTHOGONAL matrix:

$$\mathscr{T}^T = \mathscr{T}^{-1}, \tag{1.1.10a}$$

$$\sum_i \mathscr{T}_{ij}\mathscr{T}_{ik} = \delta_{jk} = \sum_i \mathscr{T}_{ji}\mathscr{T}_{ki}. \tag{1.1.10b}$$

B. Row ($n \times 1$) and Column ($1 \times n$) Matrices: State Vectors

One of Dirac's famous ideas was to associate the jth column of $\mathscr{T}(b' \leftarrow b)$ with a unit KET vector $|j\rangle$ and the jth row of $\mathscr{T}^\dagger(b' \leftarrow b)$ with a unit BRA vector $\langle j|$ as in Eq. (1.1.11), which follows. (A review of elementary vectors properties is in Appendix A.) The arrow ($\leftrightarrow$) means "is associated with."

$$j = \begin{pmatrix} \langle 1'|j\rangle \\ \langle 2'|j\rangle \\ \vdots \\ \langle n'|j\rangle \end{pmatrix} \leftrightarrow |j\rangle, \qquad j^\dagger = \overbrace{\langle j|1'\rangle\langle j|2'\rangle \cdots \langle j|n'\rangle} \leftrightarrow \langle j|. \tag{1.1.11}$$

These columns and rows of numbers represent the state j in terms of the states $1', 2', \ldots$ and n'. In the first equation (1.1.11) the numbers $j_{1'} \equiv \langle 1'|j\rangle$, $j_{2'} \equiv \langle 2'|j\rangle, \ldots, j_{n'}$, are the components of vector $|j\rangle$ in the BASIS $|1'\rangle, |2'\rangle, \ldots,$ and $|n'\rangle$, and constitute a $n \times 1$ column matrix j. Similarly $(j^\dagger)_{m'} \equiv \langle j|m'\rangle$ are components of a $1 \times n$ row matrix $j^\dagger$ that represents $\langle j|$ in the basis $\langle 1'|$, $\langle 2'|$, and $\langle n'|$.

We shall see as we go along why it is necessary to define two types of vectors $\langle|$ and $|\rangle$ that do about the same thing. However, the main point to note is that each vector $|j\rangle$ is a linear combination of any complete set of vectors $\{|1'\rangle, |2'\rangle, \ldots, |n'\rangle\}$ [Eq. (1.1.12a)] or $\{|1''\rangle, |2''\rangle, \ldots, |n''\rangle\}$, or even of the set $\{|1\rangle, |2\rangle, \ldots, |n\rangle\}$ to which it belongs, although the latter is trivially simple. These equations are Dirac's abstraction of Axiom 4 wherein the "bra" $\langle k|$ is removed from the bracket $\langle k|j\rangle$ to give KET vector $|j\rangle$:

$$|j\rangle = \sum_{i'=1}^{n} |i'\rangle\langle i'|j\rangle \equiv \sum_{i'=1}^{n} j_{i'}|i'\rangle, \tag{1.1.12a}$$

$$|j\rangle = \sum_{i''=1}^{n} |i''\rangle\langle i''|j\rangle \equiv \sum_{i''=1}^{n} j_{i''}|i''\rangle, \tag{1.1.12b}$$

$$|j\rangle = \sum_{i=1}^{n} |i\rangle\langle i|j\rangle = \sum_{i=1}^{n} \delta_{ij}|i\rangle = |j\rangle, \tag{1.1.12c}$$

For example, in the vector equation that follows we see that the state vector of spin up ($|1\rangle$) is, for small θ, a lot of spin north ($|1'\rangle$) plus a little of spin south ($|2'\rangle$):

$$\begin{aligned} |1\rangle &= |1'\rangle\langle 1'|1\rangle + |2'\rangle\langle 2'|1\rangle \\ &= \cos\frac{\theta}{2}|1'\rangle - i\sin\frac{\theta}{2}|2'\rangle \leftrightarrow \begin{pmatrix} \cos\dfrac{\theta}{2} \\ -i\sin\dfrac{\theta}{2} \end{pmatrix}. \end{aligned} \tag{1.1.13}$$

But, in its own basis it is just spin up:

$$\begin{aligned} |1\rangle &= |1\rangle\langle 1|1\rangle + |2\rangle\langle 2|1\rangle \\ &= 1|1\rangle + 0|2\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

This "mixing" of states to make other states is the embodiment of the quantum SUPERPOSITION PRINCIPLE.

The coefficients $\langle 1'|j\rangle, \langle 2'|j\rangle, \ldots$ in the column matrix j. Eq. (1.1.11), are said to be a REPRESENTATION of $|j\rangle$ in the $|1'\rangle, |2'\rangle, \ldots$ basis. Strictly

speaking, it is not correct to replace the arrow $\not{j} \leftrightarrow |j\rangle$ by an equality sign, although you will see it done in many works. When you do set $\not{j} = |j\rangle$, there is no way to tell if the stack of numbers in $\not{j}$ is supposed to be $\langle 1'|j\rangle, \langle 2'|\ j\rangle, \ldots$ or $\langle 1''|j\rangle, \langle 2''|j\rangle, \ldots$, or something else entirely. The notation $|j\rangle$ is reserved for the unique ABSTRACT VECTOR and the physical state it denotes.

One defines the SCALAR PRODUCT of a bra $\langle j|$ with a ket $|k''\rangle$ by Eq. (1.1.14a) so that it equals the bracket $\langle j|k''\rangle$. This is represented in any basis as an inner matrix product between a $(1 \times n)$ row matrix and $(n \times 1)$ column matrix, as in Eq. (1.1.14b):

$$\langle j|k''\rangle = \sum_{i'=1}^{n} \langle j|i'\rangle\langle i'|k''\rangle, \tag{1.1.14a}$$

$$\not{j} \cdot \not{k}'' = \overbrace{\langle j|1'\rangle\langle j|2'\rangle \cdots \langle j|n'\rangle} \begin{pmatrix} \langle 1'|k''\rangle \\ \langle 2'|k''\rangle \\ \vdots \\ \langle n'|k''\rangle \end{pmatrix}. \tag{1.1.14b}$$

(The dot is a common notation for the scalar or inner product. See Appendix A.)

The following definition for the transpose conjugate of vectors corresponds to the one for square matrices given previously:

$$|j\rangle^\dagger = \langle j| \leftrightarrow \begin{pmatrix} \langle 1'|j\rangle \\ \langle 2'|j\rangle \\ \vdots \\ \langle n'|j\rangle \end{pmatrix}^\dagger = \overbrace{\langle 1'|j\rangle^*\langle 2'|j\rangle^* \cdots \langle n'|j\rangle^*}$$

$$= \overbrace{\langle j|1'\rangle\langle j|2'\rangle \cdots \langle j|n'\rangle}. \tag{1.1.15}$$

Either the bra vector $\langle\Psi|$ or the ket vector $|\Psi\rangle$ defines a physical state Ψ. The scalar product requires one of each. $\langle\Psi|\Phi\rangle$ gives the amplitude for a system in state Φ to choose state Ψ. We shall always obey the axioms by requiring that $\langle\Phi|\Psi\rangle = \langle\Psi|\Phi\rangle^*$ and $\langle\Psi|\Psi\rangle = 1 = \langle\Phi|\Phi\rangle$.

The connection between transformation matrix mechanics and wave mechanics is made when each state $|i\rangle$ or $|\Psi\rangle$ is associated with a wave function

$$\langle x|i\rangle \equiv \psi_i(x) \quad \text{or} \quad \langle x|\Psi\rangle \equiv \Psi(x). \tag{1.1.16}$$

(See Section 1.5.)

C. Hamiltonian Matrices and Operators

Probably the best-known example of a matrix in quantum mechanics is the HAMILTONIAN matrix $\mathscr{H}_{ij}$. Among other things, this matrix defines the time behavior of a state $|\Psi(t)\rangle$ through the time-dependent Schrödinger equation in Eq. (1.1.17): (Planck's angular constant is $\hbar = h/2\pi = 1.05 \times 10^{-34}$ Joule second.)

$$i\hbar\frac{\partial}{\partial t}\langle i|\Psi(t)\rangle = \sum_{j=1}^{n} \mathscr{H}_{ij}\langle j|\Psi(t)\rangle, \tag{1.1.17a}$$

$$i\hbar\frac{\partial}{\partial t}\langle i|\Psi(t)\rangle = \sum_{j=1}^{n} \langle i|\mathbf{H}|j\rangle\langle j|\Psi(t)\rangle. \tag{1.1.17b}$$

The Dirac notation for matrix components $\mathscr{H}_{ij}$ is used in the second equation. The boldface symbol denotes an ABSTRACT OPERATOR **H**. In fact, the abstract Schrödinger equation is obtained [Eq. (1.1.17c)], by removing all reference to basis $\{|1\rangle, |2\rangle, \ldots, |n\rangle\}$ in Eq. (1.1.17b):

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \mathbf{H}|\Psi(t)\rangle, \tag{1.1.17c}$$

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\langle 1|\Psi(t)\rangle\\ \langle 2|\Psi(t)\rangle\\ \vdots\\ \langle n|\Psi(t)\rangle\end{pmatrix} = \begin{pmatrix}\mathscr{H}_{11} & \mathscr{H}_{12} & \cdots & \mathscr{H}_{1n}\\ \mathscr{H}_{21} & \mathscr{H}_{22} & \cdots & \mathscr{H}_{2n}\\ \vdots & & & \\ \mathscr{H}_{n1} & \mathscr{H}_{n2} & \cdots & \mathscr{H}_{nn}\end{pmatrix}\begin{pmatrix}\langle 1|\Psi(t)\rangle\\ \langle 2|\Psi(t)\rangle\\ \vdots\\ \langle n|\Psi(t)\rangle\end{pmatrix}$$

$$= \begin{pmatrix}\sum_j \mathscr{H}_{1j}\langle j|\Psi(t)\rangle\\ \sum_j \mathscr{H}_{2j}\langle j|\Psi(t)\rangle\\ \vdots\\ \sum_j \mathscr{H}_{nj}\langle j|\Psi(t)\rangle\end{pmatrix}. \tag{1.1.17d}$$

Finally, Eq. (1.1.17d) gives a matrix representation of the entire equation. It shows the inner product of the $(n \times n)$ matrix representing $\mathscr{H}$ with matrix or vector which represents $|\Psi(t)\rangle$, and the resulting new vector on the extreme right of the equation.

(a) Elementary Operators Roughly speaking, operators change old vectors into new and different ones, and matrices represent this process from some "viewpoint" or basis. We can see this clearly when we apply the abstract Axiom 4 completeness relations to the abstract operator:

$$\mathbf{H} = \sum_i \sum_j |i\rangle\langle i|\mathbf{H}|j\rangle\langle j| = \sum_i \sum_j \mathscr{H}_{ij}|i\rangle\langle j|. \tag{1.1.18}$$

Equation (1.1.18) shows clearly how $\mathscr{H}$, or any quantum operator for that matter, is a linear combination of ELEMENTARY OPERATORS $\mathbf{e}_{ij} = |i\rangle\langle j|$. (Sometimes these are called UNIT TENSOR OPERATORS.)

Each elementary operator can perform the elementary operation of Eq. (1.1.19a), converting $|j\rangle$ into $|i\rangle$, or Eq. (1.1.19b), converting bra $\langle i|$ into $\langle j|$, or Eq. (1.1.19c), zeroing all other vectors:

$$\mathbf{e}_{ij}|j\rangle = |i\rangle\langle j|j\rangle = |i\rangle, \tag{1.1.19a}$$

$$\langle i|\mathbf{e}_{ij} = \langle i|i\rangle\langle j| = \langle j|, \tag{1.1.19b}$$

$$\mathbf{e}_{ij}|k\rangle = 0 = \langle k|\mathbf{e}_{ij}, \qquad k \neq j. \tag{1.1.19c}$$

The $\mathbf{e}_{ij}$, and particularly the $\mathbf{e}_{ii}$, are sometimes called PROJECTION OPERATORS because the effect of $\mathbf{e}_{ii}$ on any vector $|\Psi\rangle$ is to yield its projection or component along the ith basis vector $|i\rangle$, while discarding the rest of $|\Psi\rangle$:

$$\mathbf{e}_{ii}|\Psi\rangle = |i\rangle\langle i|\Psi\rangle = \psi_i|i\rangle. \tag{1.1.20}$$

Note that the representation of $\mathbf{e}_{ij}$ using "its own" basis $\{|1\rangle, |2\rangle, \ldots\}$ is simply a matrix with zeros everywhere except for a 1 at the i-j position.

$$\mathbf{e}_{11} \leftrightarrow e_{11} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \qquad \mathbf{e}_{12} \leftrightarrow e_{12} = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \ldots . \tag{1.1.21}$$

The mathematical name given for the "point-to-point" combination of Dirac's vectors ($|i\rangle\langle j|$) is the OUTER PRODUCT (or sometimes TENSOR PRODUCT) of the two vectors. (Recall that "back-to-back" $\langle i|j\rangle$ is the inner product or scalar product, which is just a number). Consider a more complicated example of an operator: $|\Psi\rangle\langle\Phi|$. In the $\{|1\rangle|2\rangle \cdots |n\rangle\}$ basis we have the following representation of this object (the symbol $\otimes$ is a

notation for outer product of matrices):

$$|\Psi\rangle\langle\Phi| \leftrightarrow \begin{pmatrix} \langle 1|\Psi\rangle \\ \langle 2|\Psi\rangle \\ \vdots \\ \langle n|\Psi\rangle \end{pmatrix} \otimes \overbrace{\langle\Phi|1\rangle\langle\Phi|2\rangle \cdots \langle\Phi|n\rangle}$$

$$= \begin{pmatrix} (\langle 1|\Psi\rangle\langle\Phi|1\rangle) & (\langle 1|\Psi\rangle\langle\Phi|2\rangle) & \cdots & (\langle 1|\Psi\rangle\langle\Phi|n\rangle) \\ (\langle 2|\Psi\rangle\langle\Phi|1\rangle) & (\langle 2|\Psi\rangle\langle\Phi|2\rangle) & \cdots & (\langle 2|\Psi\rangle\langle\Phi|n\rangle) \\ \vdots & \vdots & & \\ (\langle n|\Psi\rangle\langle\Phi|1\rangle) & (\langle n|\Psi\rangle\langle\Phi|2\rangle) & \cdots & (\langle n|\Psi\rangle\langle\Phi|n\rangle) \end{pmatrix}. \tag{1.1.22}$$

Later (Chapter 6) we shall discuss products $|i\rangle|j\rangle$ and $\langle i|\langle j|$. However, by now we have seen enough to appreciate the great efficiency for notation of Dirac's bra-kets.

Nevertheless, one should be certain to keep in mind the difference between abstract quantities like $|i\rangle$, $|\Psi\rangle\langle\Phi|$, and $\mathbf{H}$ on one hand, and the representations of them on the other. The latter contain numbers which any physics must ultimately have, and in fact you will never really get a "look" at an abstract object unless you write some representation of it. However, abstract quantities are unique and correspond to some unique physical reality (supposedly), while the numbers in the representations depend upon your choice for a basis, that is, upon your viewpoint.

(b) Change of Basis It is quite common to want to change your representations from one basis $\{|1\rangle, |2\rangle, \ldots, |n\rangle\}$ to another "better" basis $\{|1'\rangle, |2'\rangle, \ldots, |n'\rangle\}$. In fact, most of this book will be concerned with the question of what is the best available basis for a given problem and how to find it.

The change can be made immediately if you know the transformation matrix $\mathscr{T}(b' \leftarrow b)$. For example, let us derive now a formula for each new component $\langle k'|\mathbf{H}|l'\rangle = \mathscr{H}_{k'l'}$ in terms of the old components $\langle i|\mathbf{H}|j\rangle = \mathscr{H}_{ij}$ and $\mathscr{T}(b' \leftarrow b)$ components. We get this by simply applying twice the Axiom 4 completeness relations for the old basis:

$$\mathscr{H}_{k'l'} \equiv \langle k'|\mathbf{H}|l'\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \langle k'|i\rangle\langle i|\mathbf{H}|j\rangle\langle j|l'\rangle, \tag{1.1.23a}$$

$$\mathscr{H}_{k'l'} = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathscr{T}_{k'i}(b' \leftarrow b)\mathscr{H}_{ij}\mathscr{T}^{\dagger}_{jl'}(b' \leftarrow b), \tag{1.1.23b}$$

$$\mathscr{H}' = \mathscr{T}(b' \leftarrow b)\mathscr{H}\mathscr{T}^{\dagger}(b' \leftarrow b). \tag{1.1.23c}$$

In Eq. (1.1.23b) we have used the definition of conjugation ($\mathscr{T}^{\dagger}_{jl'} = \mathscr{T}^{*}_{l'j}$) and Axiom 2 ($\mathscr{T}^{*}_{l'j} = \langle l'|j\rangle^{*} = \langle j|l'\rangle$).

Note that the transformation matrix $\mathscr{T}$ and its conjugate $\mathscr{T}^{\dagger}$ act left and right in Eq. (1.1.23c) in order to change the old representation $\mathscr{H}$ into the new one. Note also that the kth row of the left matrix ($\mathscr{T}$) contains the new bra vector components $\langle k'| \rightarrow \langle k'|1\rangle, \langle k'|2\rangle, \ldots$ referred to the old representation. Similarly, the lth column of the right matrix ($\mathscr{T}^{\dagger}$) contains the components of the new ket vector $|l'\rangle$.

1.2 MATRIX DIAGONALIZATION

What is the best basis to express a Hamiltonian operator? Well, the Schrödinger equation, Eq. (1.1.17), would be a lot easier to solve if we could find a basis $\{|e_1\rangle, |e_2\rangle, \ldots\}$ in which all the coupling terms $\mathscr{H}_{ij}$ vanished except for the diagonal ones, i.e., a basis that satisfied Eq. (1.2.1):

$$\langle e_i|\mathbf{H}|e_j\rangle = \delta_{ij}\varepsilon_j = \begin{cases} 0, & \text{for } i \neq j \\ \varepsilon_j, & \text{for } i = j. \end{cases} \tag{1.2.1}$$

Then Eq. (1.1.17) would reduce to n uncoupled and easily solved equations:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\langle e_1|\Psi(t)\rangle &= \varepsilon_1\langle e_1|\Psi(t)\rangle; \qquad \langle e_1|\Psi(t)\rangle = e^{-i\varepsilon_1 t/\hbar}\langle e_1|\Psi(0)\rangle, \\ i\hbar\frac{\partial}{\partial t}\langle e_2|\Psi(t)\rangle &= \varepsilon_2\langle e_2|\Psi(t)\rangle; \qquad \langle e_2|\Psi(t)\rangle = e^{-i\varepsilon_2 t/\hbar}\langle e_2|\Psi(0)\rangle \\ &\vdots \qquad\qquad\qquad\qquad\qquad\quad \vdots \end{aligned} \tag{1.2.2}$$

Now it is always possible in principle to find such a basis for any Hamiltonian, though many times it is not so easy. In fact we prove in Appendix C that it is possible to find this diagonal basis for any UNITARY matrix $\mathscr{U}$ (unitary means $\mathscr{U}^{\dagger} = \mathscr{U}^{-1}$) or for any SELF-CONJUGATE or HERMITIAN matrix $\mathscr{H}$ (Hermitian means $\mathscr{H}^{\dagger} = \mathscr{H}$). All proper Hamiltonians and all so-called "observable" operators fall into the latter category.

The state vectors $|e_i\rangle$ satisfying Eq. (1.2.1) are called EIGENVECTORS and the numbers ε_i are called EIGENVALUES. (The German heritage apparently enters into the names: "eigen" means "own.") The eigenvalues are the magnitudes of the energy quanta for the physical system. The eigenvectors describe the stationary ground and excited states.

Finding $|e_1\rangle, |e_2\rangle, \ldots$ and $\varepsilon_1, \varepsilon_2, \ldots$ given an arbitrary matrix is called the EIGENVALUE or DIAGONALIZATION PROBLEM. The abstract versions of Eqs. (1.2.1) shown in Eqs. (1.2.3a) and (1.2.3b) are called the

EIGENVALUE EQUATIONS. These are expanded in the old basis $\{|1\rangle, |2\rangle, \ldots\}$ in Eqs. (1.2.3c) and (1.2.3d) and are represented in matrix form by Eqs. (1.2.3e) and (1.2.3f). (Presumably we start off with an "old" basis and solve the equations for the "new" eigenbasis $\{|e_1\rangle, |e_2\rangle, \ldots\}$.)

$$\mathbf{H}|e_j\rangle = \varepsilon_j|e_j\rangle, \tag{1.2.3a}$$

$$\langle e_j|\mathbf{H} = \varepsilon_j\langle e_j|, \tag{1.2.3b}$$

$$\sum_m \langle l|\mathbf{H}|m\rangle\langle m|e_j\rangle = \varepsilon_j\langle l|e_j\rangle, \tag{1.2.3c}$$

$$\sum_l \langle e_j|l\rangle\langle l|\mathbf{H}|m\rangle = \varepsilon_j\langle e_j|m\rangle, \tag{1.2.3d}$$

$$\begin{pmatrix} \mathscr{H}_{11} & \mathscr{H}_{12} & \cdots & \mathscr{H}_{1n} \\ \mathscr{H}_{21} & \mathscr{H}_{22} & \cdots & \mathscr{H}_{2n} \\ \vdots & & & \vdots \\ \mathscr{H}_{n1} & \mathscr{H}_{n2} & & \mathscr{H}_{nn} \end{pmatrix} \begin{pmatrix} \langle 1|e_j\rangle \\ \langle 2|e_j\rangle \\ \vdots \\ \langle n|e_j\rangle \end{pmatrix} = \varepsilon_j \begin{pmatrix} \langle 1|e_j\rangle \\ \langle 2|e_j\rangle \\ \vdots \\ \langle n|e_j\rangle \end{pmatrix}, \tag{1.2.3e}$$

$$\overbrace{\langle e_j|1\rangle\langle e_j|2\rangle \cdots \langle e_j|n\rangle} \begin{pmatrix} \mathscr{H}_{11} & \mathscr{H}_{12} & \cdots & \mathscr{H}_{1n} \\ \mathscr{H}_{21} & \mathscr{H}_{22} & \cdots & \mathscr{H}_{2n} \\ \vdots & & & \\ \mathscr{H}_{n1} & \mathscr{H}_{n2} & \cdots & \mathscr{H}_{nn} \end{pmatrix}$$

$$= \varepsilon_j \overbrace{\langle e_j|1\rangle\langle e_j|2\rangle \cdots \langle e_j|n\rangle}. \tag{1.2.3f}$$

The equations contain several unknowns. For the right-handed eigenequations ($\mathscr{H}|e_j\rangle = \varepsilon_j|e_j\rangle$) one needs to find the ket components $\langle i|e_j\rangle$, i.e., all n^2 components of the $\mathscr{T}$ matrix, as well as n eigenvalues ε_j. The solutions to the left-handed bra equation ($\langle e_j|\mathscr{H} = \varepsilon_j\langle e_j|$) are derived easily from the kets, as we will see.

If an eigenvector's components $\langle 1|e_j\rangle, \langle 2|e_j\rangle, \ldots$ were known, we could find the corresponding eigenvalue ε_j immediately from Eq. (1.2.3c) or Eq. (1.2.3e). This, incidentally, is what symmetry theory can do, as explained in the next chapters. In the right situations it finds eigenvectors and transformations $\mathscr{T}(i \leftarrow e)$. However, for now let us suppose that they are still unknowns.

If the eigenvalues are known, there are several ways to get the eigenvectors or $\mathscr{T}(i \leftarrow e)$, and a description of these occupies most of the remainder of this review. But first let us review the equation that determines eigenvalues.

A. Obtaining Eigenvalues: The Secular Equation

Equation (1.2.3c) has been rewritten in Eq. (1.2.4a) by putting everything on the left side. The same has been done for the matrix form, Eq. (1.2.4b).

$$\sum_m \left(\langle l|\mathbf{H}|m\rangle - \delta_{ml}\varepsilon_j\right)\langle m|e_j\rangle = 0, \tag{1.2.4a}$$

$$\begin{pmatrix} \mathscr{H}_{11} - \varepsilon_j & \mathscr{H}_{12} & \cdots & \mathscr{H}_{1n} \\ \mathscr{H}_{21} & \mathscr{H}_{22} - \varepsilon_j & \cdots & \mathscr{H}_{2n} \\ \vdots & & \vdots & \\ \mathscr{H}_{n1} & \mathscr{H}_{n2} & \cdots & \mathscr{H}_{nn} - \varepsilon_j \end{pmatrix} \begin{pmatrix} \langle 1|e_j\rangle \\ \langle 2|e_j\rangle \\ \vdots \\ \langle n|e_j\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{1.2.4b}$$

We shall take an example of an eigenvalue problem involving the matrix $H = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$ and solve it along with the general case. For each equation like Eq. (1.2.4b) we will give a corresponding example indicated by Eq. $(.\ .)_x$, like

$$\begin{pmatrix} 4 - \varepsilon_j & 1 \\ 3 & 2 - \varepsilon_j \end{pmatrix} \begin{pmatrix} \langle 1|e_j\rangle \\ \langle 2|e_j\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{1.2.4b}_x$$

Note that the example could never be a Hamiltonian matrix since it is not Hermitian, but it serves as a good example for mathematical purposes.

Now Eqs. (1.2.4) can have a nonzero solution only if the determinant of the matrix vanishes. (See Appendix B.) It is the resulting equation, Eq. (1.2.5), that will determine the eigenvalues, and it is usually called the SECULAR EQUATION. Note that it is the same for bra and ket eigenequations.

$$\det \begin{vmatrix} \mathscr{H}_{11} - \varepsilon & \mathscr{H}_{12} & \cdots & \mathscr{H}_{1n} \\ \mathscr{H}_{21} & \mathscr{H}_{22} - \varepsilon & \cdots & \mathscr{H}_{2n} \\ \vdots & & & \\ \mathscr{H}_{n1} & \mathscr{H}_{n2} & & \mathscr{H}_{nn} - \varepsilon \end{vmatrix}$$

$$\equiv S(\varepsilon) \equiv \varepsilon^n + a_1\varepsilon^{n-1} + \cdots + a_{n-1}\varepsilon + a_n = 0. \tag{1.2.5}$$

$$\det \begin{vmatrix} 4 - \varepsilon & 1 \\ 3 & 2 - \varepsilon \end{vmatrix} \equiv S(\varepsilon) = \varepsilon^2 - 6\varepsilon + 5 = 0. \tag{1.2.5}_x$$

It can be a tedious job to solve an nth-degree secular equation. It is easy enough for small matrices like our example, for which the roots are $(\varepsilon_1 = 1)$ and $(\varepsilon_2 = 5)$. But one should know that no formula exists for roots of quintic,

sextic, or higher-degree equations. All nth-degree polynomial equations have exactly n roots somewhere in the complex plane, but no formula exists to give them all for $n \geq 5$ (Abels's theorem). In general finding the eigenvalues ε_j can be the hardest part of the problem, if the eigenvectors are not known.

B. Obtaining Eigenvectors: The Hamilton-Cayley Equation

One very helpful fact is that any matrix satisfies its own secular equation. This is the content of the following theorem. The resulting equation, Eq. (1.2.6), is called the HAMILTON-CAYLEY EQUATION (HCEq).

Hamilton-Cayley Theorem:: If in the secular equation

$$S(\varepsilon) = \det|\mathscr{H} - \varepsilon\mathbf{1}| = \varepsilon^n + a_1\varepsilon^{n-1} + \cdots + a_{n-1}\varepsilon + a_n = 0$$

one replaces the terms $a_m\varepsilon^{n-m}$ by the matrices $a_m\mathscr{H}^{n-m}$, the constant term a_n by the matrix $a_n\mathbf{1}$, and the 0 by the zero matrix $\mathbf{0}$, then the resulting matrix equation, Eq. (1.2.6), is valid:

$$S(\mathscr{H}) = \mathscr{H}^n + a_1\mathscr{H}^{n-1} + \cdots + a_{n-1}\mathscr{H} + a_n\mathbf{1} = \mathbf{0}, \tag{1.2.6}$$

$$S\left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}\right] = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}^2 - 6\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} + 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 19 & 6 \\ 18 & 7 \end{pmatrix} - \begin{pmatrix} 24 & 6 \\ 18 & 12 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{1.2.6}_x$$

We can prove this theorem by using the fact that the product of a matrix $\mathscr{M}$ with its ADJUNCT matrix $\mathscr{M}^{\text{adj}}$ equals $\det\mathscr{M}$ times a unit matrix. [As explained in Appendix B, $\mathscr{M}^{\text{adj}}_{ij}$ is $(-1)^{i+j}$ multiplied by the determinant of $\mathscr{M}$ taken with its jth row and ith column missing.] This gives the following:

$$\sum_j \mathscr{M}^{\text{adj}}_{ij}\mathscr{M}_{jk} = \sum_j \mathscr{M}_{ij}\mathscr{M}^{\text{adj}}_{jk} = \delta_{ik}(\det\mathscr{M}).$$

Setting $\mathscr{M} = \mathscr{H} - \varepsilon\mathbf{1}$ we have

$$\mathscr{M}^{\text{adj}}(\mathscr{H} - \varepsilon\mathbf{1}) = \det(\mathscr{H} - \varepsilon\mathbf{1})\mathbf{1} = (\mathscr{H} - \varepsilon\mathbf{1})\mathscr{M}^{\text{adj}},$$

$$\mathscr{M}^{\text{adj}}(\mathscr{H} - \varepsilon\mathbf{1}) = S(\varepsilon)\mathbf{1} = (\mathscr{H} - \varepsilon\mathbf{1})\mathscr{M}^{\text{adj}}. \tag{1.2.7}$$

Now replacement of the number ε by the matrix $\mathscr{H}$ is well defined, and doing so proves the theorem: $S(\mathscr{H}) = \mathbf{0}$.

(a) Eigenvector Projectors If the roots ε_j are known, we may use the factored form of the Eq. (1.2.6) to derive the eigenvectors:

$$(\mathscr{H} - \varepsilon_1\mathbf{1})(\mathscr{H} - \varepsilon_2\mathbf{1}) \cdots (\mathscr{H} - \varepsilon_n\mathbf{1}) = \mathbf{0}, \tag{1.2.8}$$

$$\left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} - 1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} - 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{1.2.8}_x$$

The procedure depends upon whether or not any of the n roots ε_j are equal. We study first the cases in which all ε_j are distinct. This includes our example ($\varepsilon_1 = 1$, $\varepsilon_2 = 5$).

Case i: All n Roots ε_i Are Distinct

With all roots having different values one may select n RELATIVELY PRIME factors $\mathscr{P}_j$ ("relatively prime" means "no common factor shared by all") from the factored HCEq, Eq. (1.2.8), by deleting first one $(\mathscr{H} - \varepsilon_j\mathbf{1})$ factor, then the next, and so on:

$$\begin{aligned}
\mathscr{P}_1 &= \mathbf{1} \cdot (\mathscr{H} - \varepsilon_2\mathbf{1}) \cdot \qquad (\cdots) \cdot \quad (\mathscr{H} - \varepsilon_n\mathbf{1}), \\
\mathscr{P}_2 &= (\mathscr{H} - \varepsilon_1\mathbf{1}) \cdot \quad \mathbf{1} \cdot \qquad (\cdots) \cdot \quad (\mathscr{H} - \varepsilon_n\mathbf{1}), \\
&\vdots \\
\mathscr{P}_n &= (\mathscr{H} - \varepsilon_1\mathbf{1}) \cdot (\mathscr{H} - \varepsilon_2\mathbf{1}) \cdot (\cdots) \cdot \quad \mathbf{1}.
\end{aligned} \tag{1.2.9}$$

$$\mathscr{P}_1 = \mathbf{1} \cdot \left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} - 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = \begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix},$$

$$\mathscr{P}_2 = \left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} - 1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \cdot \mathbf{1} = \begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix}. \tag{1.2.9}_x$$

Now we see that these $\mathscr{P}_j$ contain the eigenvectors, if we rewrite the HCEq, Eq. (1.2.8), in terms of $\mathscr{P}_j$ as is done in the following. Equations (1.2.10b) and (1.2.10c) are essentially the eigenvalue equations (1.2.3e) and (1.2.3f) with matrix $\mathscr{P}_j$ in place of the eigenvector:

$$(\mathscr{H} - \varepsilon_j\mathbf{1})\,\mathscr{P}_j = \mathbf{0}, \tag{1.2.10a}$$

$$\mathscr{H}\mathscr{P}_j = \varepsilon_j\mathscr{P}_j, \tag{1.2.10b}$$

$$\mathscr{P}_j\mathscr{H} = \varepsilon_j\mathscr{P}_j. \tag{1.2.10c}$$

According to the last equations, each column of $\mathscr{P}_j$ must satisfy the eigenvalue equation for a ket vector $|e_j\rangle$, and each row must satisfy this equation

for a bra vector $\langle e_j|$. In Eq. $(1.2.10)_x$ we write only the first columns and rows of $\mathscr{P}_j$:

$$\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} -1 & \\ 3 & \end{pmatrix} = 1\begin{pmatrix} -1 & \\ 3 & \end{pmatrix},$$

$$\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 3 & \\ 3 & \end{pmatrix} = 5\begin{pmatrix} 3 & \\ 3 & \end{pmatrix},$$

$$\begin{pmatrix} -1 & 1 \\ & \end{pmatrix}\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} = 1\begin{pmatrix} -1 & 1 \\ & \end{pmatrix},$$

$$\begin{pmatrix} 3 & 1 \\ & \end{pmatrix}\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} = 5\begin{pmatrix} 3 & 1 \\ & \end{pmatrix}. \qquad (1.2.10b)_x$$

Now if a vector satisfies an eigenvalue equation, so will two times that vector, or any "length" of that vector including zero. It is conventional to normalize all our eigenvectors to give $\langle e_j|e_j\rangle = 1$.

This normalization can be done if we divide each $\mathscr{P}_j$ by the right factor to make it into an IDEMPOTENT matrix $\mathscr{P}_j$. (Something is called idempotent if it will give itself back when multiplied by itself. **1** is an example: $\mathbf{1} \cdot \mathbf{1} = \mathbf{1}$.) The following equation follows from the definition (1.2.9) of $\mathscr{P}_i$, i.e., all $(\mathscr{H} - \varepsilon_l \mathbf{1})$ factors except the ith:

$$\mathscr{P}_i \mathscr{P}_j = \prod_{l \neq i} (\mathscr{H} - \varepsilon_l \mathbf{1})\, \mathscr{P}_j = \prod_{l \neq j} (\varepsilon_j - \varepsilon_l)\, \mathscr{P}_j, \qquad \text{for } i = j$$

$$= \mathbf{0}, \qquad \text{for } i \neq j. \quad (1.2.11)$$

This implies that the $\mathscr{P}_j$ defined by Eq. (1.2.12) are idempotent $[\mathscr{P}_j^2 = \mathscr{P}_j]$. Note also that different $\mathscr{P}_j$ or $\mathscr{P}_j$ will be orthogonal ($\mathscr{P}_i \mathscr{P}_j = 0$ for $i \neq j$):

$$\mathscr{P}_j \equiv \frac{\prod_{l \neq j} (\mathscr{H} - \varepsilon_l \mathbf{1})}{\prod_{l \neq j} (\varepsilon_j - \varepsilon_l)}, \qquad (1.2.12)$$

$$\mathscr{P}_1 = \frac{\begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix}}{1 - 5} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix}, \qquad \mathscr{P}_2 = \frac{\begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix}}{5 - 1} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}. \qquad (1.2.12)_x$$

Now, because $\mathscr{P}_j \mathscr{P}_j = \mathscr{P}_j$, the desired normalization factor N_j is found on the diagonal of $\mathscr{P}_j$ at the intersection of the row and column that were picked to be the eigenbra $\langle e_j|$ and eigenket $|e_j\rangle$, respectively. The vectors given in Eq.

(1.2.13) satisfy the normalization condition $\langle e_i|e_j\rangle = \delta_{ij}$.

$$\langle e_j| \leftrightarrow \frac{\overbrace{\text{chosen row}}}{\sqrt{N_j}}, \qquad |e_j\rangle \leftrightarrow \frac{1}{\sqrt{N_j}} \begin{pmatrix} \text{c} \\ \text{h} \\ \text{o} \\ \text{s} \\ \text{e} \\ \text{n} \end{pmatrix} \begin{matrix} \text{c} \\ \text{o} \\ \text{l} \\ \text{u} \\ \text{m} \\ \text{n} \end{matrix}, \tag{1.2.13}$$

$$\langle e_1| \leftrightarrow \frac{\overbrace{\frac{1}{4} \quad -\frac{1}{4}}}{\sqrt{\frac{1}{4}}}, \qquad |e_1\rangle \leftrightarrow \frac{1}{\sqrt{\frac{1}{4}}} \begin{pmatrix} \frac{1}{4} \\ -\frac{3}{4} \end{pmatrix}$$

$$\langle e_2| \leftrightarrow \frac{\overbrace{\frac{3}{4} \quad \frac{1}{4}}}{\sqrt{\frac{3}{4}}}, \qquad |e_2\rangle \leftrightarrow \frac{1}{\sqrt{\frac{3}{4}}} \begin{pmatrix} \frac{3}{4} \\ \frac{3}{4} \end{pmatrix}. \tag{1.2.13}_{\text{x}}$$

According to the convention just given, the normalization is divided equally between the left (bra) and right (ket) eigenvectors. This allows the left and right eigenvectors of a Hermitian (or unitary) operator to be related by transpose conjugation ($\langle e_i| = |e_i\rangle^\dagger$) as proved in Problem 1.2.7. However, this relation does not hold between left and right eigenvectors of a non-Hermitian matrix such as the (2 × 2) matrix, which is our example. In fact, its bra and ket vectors point in quite different "directions" as seen in Figure 1.2.1. (Compare $|e_j\rangle$ with $\langle e_j|$.) Nevertheless, $\langle e_1|$ is perpendicular to $|e_2\rangle$ and $\langle e_2|$ is perpendicular to $|e_1\rangle$, as guaranteed by Eq. (1.2.11). The figure gives a picture of a DUAL SPACE consisting of left or bra eigenvectors $\langle\ |$ on one hand, and the "reciprocal space" of right or ket vectors $|\ \rangle$ on the other.

Finally, notice that the $\mathscr{P}_j$ matrices are just representations of the PROJECTION OPERATORS $|e_j\rangle\langle e_j|$ discussed around Eq. (1.1.20):

$$\mathbf{e}_{e_1e_1} = |e_1\rangle\langle e_1| \leftrightarrow \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{2} \end{pmatrix} \otimes \underbrace{\overbrace{\frac{1}{2} \quad -\frac{1}{2}}} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} = \mathscr{P}_1,$$

$$\mathbf{e}_{e_2e_2} = |e_2\rangle\langle e_2| \leftrightarrow \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \otimes \underbrace{\overbrace{\frac{\sqrt{3}}{2} \frac{1}{2\sqrt{3}}}} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} = \mathscr{P}_2. \tag{1.2.14}$$

Case ii: Some Roots ε_j are Equal

If a matrix has some repeated roots in its secular equation, then it is possible that the matrix itself will satisfy a polynomial equation of lower degree than the HCEq. The lowest degree equation satisfied by any matrix or operator is called the MINIMAL EQUATION (MEq).

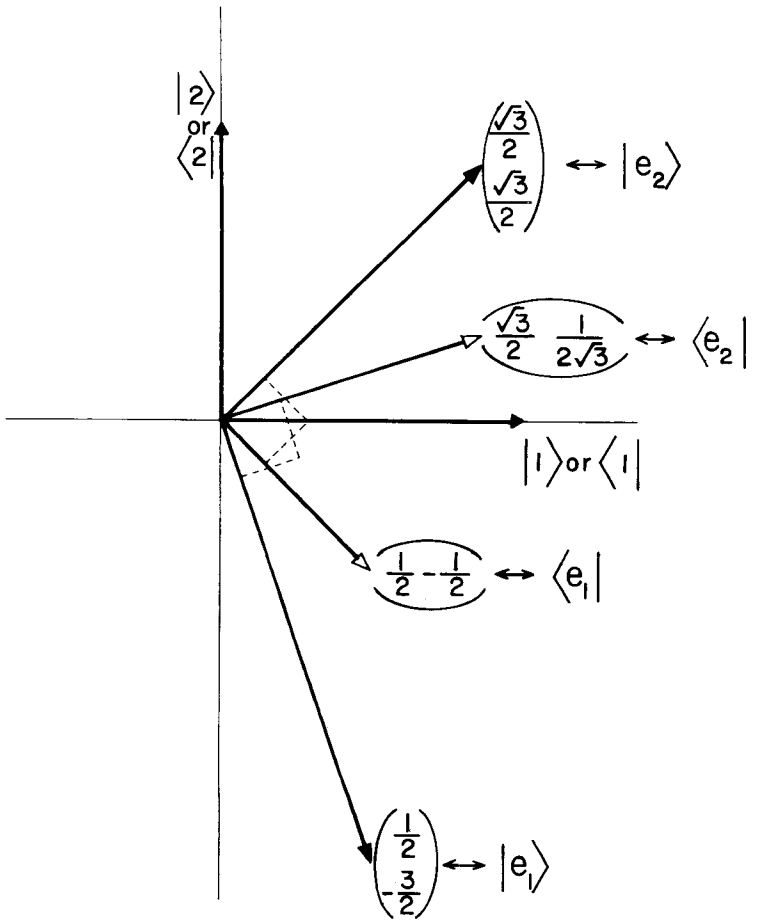


Figure 1.2.1 Left ($\langle e_j|$) and right ($|e_j\rangle$) eigenvectors of the matrix $\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$.

If the MEq has repeated roots, too, then the matrix is not diagonalizable. It can still be brought to an almost diagonal Jordan form, but we shall not need to discuss this. Most matrices which will be discussed in this text are Hermitian or unitary, and they are therefore diagonalizable. (See Appendix C.) The roots of their MEq are distinct even if there are any number of repeated roots in the HCEq.

For all such cases the theory from Eq. (1.2.8) to Eq. (1.2.12) applies except we deal with the MEq instead of the HCEq. The projection matrix defined in the following obeys the same rules ($\mathscr{P}_i\mathscr{P}_j = \delta_{ij}\mathscr{P}_i$):

$$\mathscr{P}_{\varepsilon_j} = \frac{\prod_{\varepsilon_l \neq \varepsilon_j} (\mathscr{H} - \varepsilon_l \mathbf{1})}{\prod_{\varepsilon_l \neq \varepsilon_j} (\varepsilon_j - \varepsilon_l)}. \tag{1.2.15}$$

For example, the matrix

$$\mathscr{H} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

has a fourth-degree HCEq $[(H - \mathbf{1})^2(H + \mathbf{1})^2 = \mathbf{0}]$ but only a second-degree MEq $[(H - \mathbf{1})(H + \mathbf{1}) = \mathbf{0}]$. Hence, only two idempotents need to be constructed. These are the following:

$$\mathscr{P}_1 = \frac{\mathscr{H} - (-1)\mathbf{1}}{1 - (-1)} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix},$$

$$\mathscr{P}_{-1} = \frac{\mathscr{H} - \mathbf{1}}{(-1 - 1)} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}. \tag{1.2.16}$$

However, these idempotents contain more than one independent eigenvector apiece. $\mathscr{P}_1$ has two: its first row: $\langle e_1| \rightarrow \overbrace{1/\sqrt{2} \quad 0 \quad 0 \quad 1/\sqrt{2}}$ and its second row: $\langle e_1'| \rightarrow \overbrace{0 \quad 1/\sqrt{2} \quad 1/\sqrt{2} \quad 0}$. $\mathscr{P}_{-1}$ has two also. In general, a $\mathscr{P}_{\varepsilon_j}$ will have a number of independent eigenvectors equal to the number of times ε_j is repeated in the HCEq.

Any linear combination of rows (or of columns) of a $\mathscr{P}_{\varepsilon_j}$ matrix will be an eigenbra (eigenket) vector, so these are not uniquely defined when ε_j is repeated. Furthermore, the vectors $|e_j^\alpha\rangle$ from row α may not be orthogonal to $|e_j^\beta\rangle$ from column β likewise for columns. However, the same trick works again: $\langle e_j^\alpha | e_j^\beta \rangle$ is the component $(\mathscr{P}_{\varepsilon_j})_{\alpha\beta}$ of the idempotent projector. So Gram-Schmidt orthogonalization and normalization can be done directly.

(b) Completeness and Orthogonality of Projectors The orthonormality relation, Eq. (1.2.17), and its consequences were demonstrated in the preceding section:

$$\mathscr{P}_{\varepsilon_i}\mathscr{P}_{\varepsilon_j} = \delta_{\varepsilon_i\varepsilon_j}\mathscr{P}_{\varepsilon_i}. \tag{1.2.17}$$

What needs to be shown now is the completeness relation, Eq. (1.2.18), for

these same projection operators:

$$\mathscr{P}_{\varepsilon_i} + \mathscr{P}_{\varepsilon_j} + \cdots + \mathscr{P}_{\varepsilon_n} = \mathbf{1} \qquad \text{(sum over disctinct } \varepsilon_j\text{)}. \tag{1.2.18}$$

$$\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix} + \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.2.18}_{\mathrm{x}}$$

Proving Eq. (1.2.18) abstractly is easy for most mathematicians, if not for most physicists. Most mathematics students have encountered the same idea repeatedly in number theory, then in algebra, and so forth. In number theory or algebra one proves that integers or polynomials $\{p_1 p_2 \cdots p_n\}$ are relatively prime if and only if there exists some integers or polynomials $\{a_1 a_2 \cdots a_n\}$ that $\sum_i a_i p_i = 1$. The same proof works for the relatively prime matrix polynomials $\mathscr{P}_i$ defined in Eq. (1.2.9), or for the normalized $\mathscr{P}_j$ matrices:

$$\sum \alpha_j \mathscr{P}_j = \mathbf{1}.$$

Now operation with $\mathscr{P}_j$ gives $\alpha_j \mathscr{P}_j = \mathscr{P}_j$ and Eq. (1.2.18) is proved.

A less abstract proof follows from elementary algebra. Note that for any two numbers ε_1 and ε_2 we have that

$$\frac{m - \varepsilon_2}{\varepsilon_1 - \varepsilon_2} + \frac{m - \varepsilon_1}{\varepsilon_2 - \varepsilon_1} = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 - \varepsilon_2} = 1.$$

For any three numbers ε_1, ε_2, and ε_3 we have that

$$\frac{(m - \varepsilon_2)(m - \varepsilon_3)}{(\varepsilon_1 - \varepsilon_2)(\varepsilon_1 - \varepsilon_3)} + \frac{(m - \varepsilon_1)(m - \varepsilon_3)}{(\varepsilon_2 - \varepsilon_1)(\varepsilon_2 - \varepsilon_3)}$$

$$+ \frac{(m - \varepsilon_1)(m - \varepsilon_2)}{(\varepsilon_3 - \varepsilon_1)(\varepsilon_3 - \varepsilon_2)} = \frac{(\varepsilon_1 - \varepsilon_2)(\varepsilon_1 - \varepsilon_3)(\varepsilon_2 - \varepsilon_3)}{(\varepsilon_1 - \varepsilon_2)(\varepsilon_1 - \varepsilon_3)(\varepsilon_2 - \varepsilon_3)} = 1,$$

and so forth. In other words, one may substitute arbitrary values for the ε_j in Eqs. (1.2.12) and (1.2.15) for the $\mathscr{P}_j$ and still have them satisfy the completeness relation in Eq. (1.2.18)! (On the other hand, we have seen that the orthonormality relations hold only if the ε_j are the eigenvalues of matrix $\mathscr{H}$.)

It is interesting to note a similarity between the form of the completeness relation for projection operators of matrices and the Lagrangian interpolation formula which is used in numerical analysis. Suppose a function $f(x)$ of a real variable has n values $f(x_1), f(x_2), \ldots, f(x_n)$, which are known at n distinct points $x_1, x_2, \ldots, x_n$. Then the following $(n-1)$th-degree polyno-

mial $L(x)$ is constructed to exactly equal $f(x)$ at each point x_j.

$$L(x) = \sum_{j=1}^{n} f(x_j) \frac{\prod_{k\neq j}^{n}(x - x_k)}{\prod_{k\neq j}^{n}(x_j - x_k)}.$$

The coefficient of $f(x_j)$ in the preceding expression for $L(x)$ should remind you of the formula for P_j given by Eq. (1.2.12) if x is replaced by an n-dimensional matrix and x_j are its eigenvalues.

Clearly, $L(x) = f(x)$ at each point $x = x_j$. How well $L(x)$ represents $f(x)$ for $x \neq x_j$ depends upon how much $f(x)$ deviates from an $(n-1)$th degree polynomial in between these points. Simple functions such as $f(x) = 1$ or $f(x) = x$ have the following polynomial representations for $n > 1$:

$$1 = \sum_{j=1}^{n} \frac{\prod_{k\neq j}^{n}(x - x_k)}{\prod_{k\neq j}^{n}(x_j - x_k)}, \qquad x = \sum_{j=1}^{n} x_j \frac{\prod_{k\neq j}^{n}(x - x_k)}{\prod_{k\neq j}^{n}(x_j - x_k)}.$$

These have precisely the same form as the completeness relation (1.2.18) and the spectral decomposition relation which is discussed in the following. [See Eq. (1.2.21).]

Finally, we observe that, while Eqs. (1.2.17) and (1.2.18) appear very dissimilar, they are more similar if you represent them with Dirac's notation. The completeness relation for basis $|e_j\rangle$ is represented by the following in the basis $\{|1\rangle, |2\rangle, \ldots\}$:

$$\mathbf{1} = \sum_j |e_j\rangle\langle e_j|, \tag{1.2.19a}$$

$$\langle l|\mathbf{1}|m\rangle = \delta_{lm} = \sum_j \langle l|e_j\rangle\langle e_j|m\rangle. \tag{1.2.19b}$$

The orthonormality condition for basis $|e_j\rangle$ is similarly represented:

$$\langle e_i|e_j\rangle = \delta_{ij} = \sum_l \langle e_i|l\rangle\langle l|e_j\rangle. \tag{1.2.20}$$

The similarity between Eqs. (1.2.20) and (1.2.19) is now more evident.

(c) Spectral Decomposition of Matrices Any diagonalizable matrix $\mathscr{H}$ will have a set of idempotent projection matrices $\mathscr{P}_k$ satisfying the orthonormality [Eq. (1.2.17)] and completeness [Eq. (1.2.18)] relations. Now operating on Eq. (1.2.18) with the matrix $\mathscr{H}$ gives Eq. (1.2.21), which is called the

SPECTRAL DECOMPOSITION of $\mathscr{H}$ [eigenvalue equations (1.2.10) are used]:

$$\mathscr{H} = \varepsilon_i \mathscr{P}_{\varepsilon_i} + \varepsilon_j \mathscr{P}_{\varepsilon_j} + \cdots + \varepsilon_n \mathscr{P}_{\varepsilon_n} \qquad \text{(sum over distinct } \varepsilon_j\text{)}, \quad (1.2.21)$$

$$\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} + 5 \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}. \qquad (1.2.21)_{\mathrm{x}}$$

This decomposition provides a very elegant way to manipulate a matrix. For example, if you wanted the 50th power of the matrix $\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$ there is no need to begin multiplying it by itself 49 times. Instead, the spectral decomposition gives the answer quickly:

$$\begin{aligned} \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}^{50} &= (\varepsilon_1 \mathscr{P}_1 + \varepsilon_2 \mathscr{P}_2)^{50} \\ &= \varepsilon_1^{50} \mathscr{P}_1^{50} + 50 \varepsilon_1^{49} \mathscr{P}_1^{49} \varepsilon_2 \mathscr{P}_2 + \cdots + \varepsilon_2^{50} \mathscr{P}_2^{50} \\ &= \varepsilon_1^{50} \mathscr{P}_1 + \varepsilon_2^{50} \mathscr{P}_2 \\ &= 1 \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} + 5^{50} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} \\ &= \tfrac{1}{4} \begin{pmatrix} 1 + 3 \cdot 5^{50} & 5^{50} - 1 \\ 3 \cdot 5^{50} - 3 & 3 + 5^{50} \end{pmatrix}. \end{aligned} \qquad (1.2.22)$$

The property $\mathscr{P}_i \mathscr{P}_j = \delta_{ij} \mathscr{P}_j$ reduces a complicated expression to a very simple one. It is this type of analysis that is central to the development of symmetry analysis and its applications, which are given in the next chapters.

(d) Simultaneous Diagonalization of Commuting Matrices When any two Hermitian matrices $\mathscr{M}$ and $\mathscr{N}$ commute with each other $[\mathscr{M}\mathscr{N} = \mathscr{N}\mathscr{M}]$ then there must exist a complete set of base vectors that are eigenvectors of both matrices. This is constructed easily using the spectral decompositions and completeness relations for each matrix. These are shown in Eqs. (1.2.23), where it is assumed that the number of distinct eigenvalues of $\mathscr{M}$ is m, while for $\mathscr{N}$ the number is n:

$$\mathbf{1} = \mathscr{P}^1(\mathscr{M}) + \mathscr{P}^2(\mathscr{M}) + \cdots + \mathscr{P}^m(\mathscr{M}), \qquad (1.2.23a)$$

$$\mathscr{M} = \mu_1 \mathscr{P}^1(\mathscr{M}) + \mu_2 \mathscr{P}^2(\mathscr{M}) + \cdots + \mu_m \mathscr{P}^m(\mathscr{M}), \qquad (1.2.23b)$$

$$\mathbf{1} = \mathscr{P}^1(\mathscr{N}) + \mathscr{P}^2(\mathscr{N}) + \cdots + \mathscr{P}^n(\mathscr{N}), \qquad (1.2.23c)$$

$$\mathscr{N} = \nu_1 \mathscr{P}^1(\mathscr{N}) + \nu_2 \mathscr{P}^2(\mathscr{N}) + \cdots + \nu_n \mathscr{P}^n(\mathscr{N}). \qquad (1.2.23d)$$

Now if the matrices commute, so will their respective idempotents, since these are just polynomials of the matrices. Because of this, we get immediately a complete set of idempotents and hence the eigenvectors for the two of them by simply multiplying the separate completeness relations:

$$\mathbf{1} = \mathbf{1}\cdot\mathbf{1} = \left[\mathscr{P}^1(\mathscr{M}) + \mathscr{P}^2(\mathscr{M}) + \cdots + \mathscr{P}^m(\mathscr{M})\right] \times \left[\mathscr{P}^1(\mathscr{N}) + \mathscr{P}^2(\mathscr{N}) + \cdots + \mathscr{P}^n(\mathscr{N})\right], \tag{1.2.24a}$$

$$\mathbf{1} = \mathscr{P}^1(\mathscr{M})\mathscr{P}^1(\mathscr{N}) + \mathscr{P}^1(\mathscr{M})\mathscr{P}^2(\mathscr{N}) + \cdots + \mathscr{P}^2(\mathscr{M})\mathscr{P}^1(\mathscr{N}) + \cdots, \tag{1.2.24b}$$

$$\mathbf{1} = \mathscr{P}^{1,1} + \mathscr{P}^{1,2} \cdots + \cdots + \mathscr{P}^{2,1} + \cdots + \cdots . \tag{1.2.24c}$$

Each of the resulting terms must satisfy eigenvalue equations for $\mathscr{M}$ and $\mathscr{N}$:

$$\mathscr{M}\mathscr{P}^{\alpha,\beta} = \mathscr{M}\mathscr{P}^{\alpha}(\mathscr{M})\mathscr{P}^{\beta}(\mathscr{N}) = \mu_\alpha \mathscr{P}^{\alpha,\beta},$$
$$\mathscr{N}\mathscr{P}^{\alpha,\beta} = \mathscr{N}\mathscr{P}^{\alpha}(\mathscr{M})\mathscr{P}^{\beta}(\mathscr{N}) = \mathscr{N}\mathscr{P}^{\beta}(\mathscr{N})\mathscr{P}^{\alpha}(\mathscr{M}) = \nu_\beta \mathscr{P}^{\alpha,\beta}.$$

Each term $\mathscr{P}^{\alpha,\beta}$ that is not zero is clearly one of a set of complete and orthogonal idempotents. However, many of the terms are usually zero in practice.

In fact, if $\mathscr{M}$ is an $m \times m$ matrix with m distinct eigenvalues, then there will be exactly m nonzero terms on the right of (1.2.24b) or (1.2.24c), and each of these will be identical to one of the $\mathscr{P}^{\alpha}(\mathscr{M})$. Having two or more nonzero terms $\mathscr{P}^{\alpha,\beta}, \mathscr{P}^{\alpha,\beta'}, \ldots$ amounts to having two orthogonal eigenvectors with eigenvalue μ_α where we said we only had one.

In other words, the reduction of $\mathscr{M}$ will, in the case of distinct μ_α, serve immediately to reduce all the other matrices $\mathscr{N}$ which commute with $\mathscr{M}$. For example, the fact that matrix $\mathscr{N} = \begin{pmatrix} 5 & -1 \\ -3 & 7 \end{pmatrix}$ commutes with the matrix $\mathscr{M} = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$, which we have reduced, implies that $\mathscr{N}$ is reduced by the same idempotents. Solving its secular equation is unnecessary:

$$\mathscr{M} = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} = 1\cdot\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} + 5\cdot\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix},$$
$$\mathscr{N} = \begin{pmatrix} 5 & -1 \\ -3 & 7 \end{pmatrix} = 8\cdot\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} + 4\cdot\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}.$$

This idea is extremely important in the development of symmetry analysis, as we shall see early in the next chapter.

(e) Obtaining Eigenvectors from Adjuncts To find the eigenvectors $|e_j\rangle$ or $\langle e_j|$ using the projection matrices one needed all eigenvalues

$\varepsilon_1\varepsilon_2 \cdots \varepsilon_n$ except ε_j. A competitive method exists which gives the vectors corresponding to a certain eigenvalue ε_j when only that ε_j is known.

We rewrite Eq. (1.2.7) in the following and set ε equal to the eigenvalue ε_j:

$$\mathscr{M}^{\text{adj}}(\mathscr{H} - \varepsilon_j\mathbf{1}) = S(\varepsilon_j)\mathbf{1} = (\mathscr{H} - \varepsilon_j\mathbf{1})\mathscr{M}^{\text{adj}}.$$

Then the secular equation must hold [$S(\varepsilon_j) = 0$] and we are left with Eqs. (1.2.25):

$$\mathscr{M}^{\text{adj}}(\mathscr{H} - \varepsilon_j\mathbf{1}) = \mathbf{0} = (\mathscr{H} - \varepsilon_j\mathbf{1})\mathscr{M}^{\text{adj}}, \tag{1.2.25a}$$

$$\mathscr{H}\mathscr{M}^{\text{adj}} = \varepsilon_j\mathscr{M}^{\text{adj}}, \tag{1.2.25b}$$

$$\mathscr{M}^{\text{adj}}\mathscr{H} = \varepsilon_j\mathscr{M}^{\text{adj}}. \tag{1.2.25c}$$

We see that the rows and columns of the adjunct matrix $\mathscr{M}^{\text{adj}}$ derived from $(\mathscr{H} - \varepsilon_j\mathbf{1}) = \mathscr{M}$ must satisfy the eigenvalue equations of bras and kets, respectively, according to Eqs. (1.2.25b) and (1.2.25c).

$$\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix} = 1\begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix},$$

$$\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} -3 & -1 \\ -3 & -1 \end{pmatrix} = 5\begin{pmatrix} -3 & -1 \\ -3 & -1 \end{pmatrix}, \tag{1.2.25b}_x$$

$$\begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} = 1\begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix}$$

$$\begin{pmatrix} -3 & -1 \\ -3 & -1 \end{pmatrix}\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} = 5\begin{pmatrix} -3 & -1 \\ -3 & -1 \end{pmatrix}. \tag{1.2.25c}_x$$

Comparison of the resulting vectors (derived by inspecting the examples in Eq. $(1.2.15b)_x$ and $(1.2.15c)_x$ with the previously derived ones [Eq. $(1.2.13)_x$ or Figure 1.2.1] shows that they differ only by an overall factor. (Normalization is still needed.)

1.3 SOME OTHER MATRICES USED IN QUANTUM MECHANICS

Different types of quantum mechanical problems can be made more convenient to solve by using various matrix expressions. The mathematics is the same but the physical meaning can be quite different. We shall come to deal with the following examples of matrices and operators in later chapters.

A. Scattering Matrix

Imagine that free electrons can approach or leave some central "scattering" region in which they are not entirely free, and suppose that they can use n different paths or channels to go in and out. Their wave functions $\psi_l(r)$ in each channel l where they are free can be written as a linear combination of an outgoing part and an ingoing part with coefficients A_l^o and A_l^I, respectively:

$$\Psi_l(r) = A_l^o e^{ikr} + A_l^I e^{-ikr}.$$

Now the boundary conditions associated with the junctions of each channel with the scattering region will, in general, give a set of linear relations between these coefficients. The most well-known relation involves the *S*-MATRIX in Eq. (1.3.1) and expresses outgoing coefficients A_l^o in terms of ingoing ones A_l^I.

$$\begin{pmatrix} A_1^o \\ A_2^o \\ \vdots \\ A_n^o \end{pmatrix} = \begin{pmatrix} \mathscr{S}_{11} & \mathscr{S}_{12} & \cdots & \mathscr{S}_{1n} \\ \mathscr{S}_{21} & \mathscr{S}_{22} & \cdots & \mathscr{S}_{2n} \\ \vdots & \vdots & & \vdots \\ \mathscr{S}_{n1} & \mathscr{S}_{n2} & & \mathscr{S}_{nn} \end{pmatrix} \begin{pmatrix} A_1^I \\ A_2^I \\ \vdots \\ A_n^I \end{pmatrix}. \tag{1.3.1}$$

Generally we require conservation of probability, Eq. (1.3.2), in such problems; that is, we make sure that no particles go in and get lost in the scattering region:

$$\sum_{l=1}^{n} A_l^{I*} A_l^I = \sum_{j=1}^{n} A_j^{o*} A_j^o. \tag{1.3.2}$$

Expanding Eq. (1.3.2) using Eq. (1.3.1) we obtain

$$\sum_{l=1}^{n} A_l^{I*} A_l^I = \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} \mathscr{S}_{jk}^* A_K^{I*} \mathscr{S}_{jl} A_l^I = \sum_{k=1}^{n} \sum_{l=1}^{n} \sum_{j=1}^{n} \mathscr{S}_{jk}^* \mathscr{S}_{jl} A_k^{I*} A_l^I. \tag{1.3.3}$$

This must be true for all choices of A_i^I. This yields Eq. (1.3.4), i.e., it shows $\mathscr{S}$ must be UNITARY:

$$\sum_{j=1}^{n} \mathscr{S}_{jk}^* \mathscr{S}_{jl} = \delta_{kl}, \tag{1.3.4a}$$

$$\sum_{j=1}^{n} \mathscr{S}_{kj}^\dagger \mathscr{S}_{jl} = \delta_{kl}, \tag{1.3.4b}$$

$$\mathscr{S}^\dagger \mathscr{S} = 1. \tag{1.3.4c}$$

As in the case of the Hamiltonian. there is considerable advantage to using the eigenvectors of $\mathscr{S}$. This is especially true when symmetry analysis can provide them, as we will discuss in the following chapter.

B. Density Matrix

So far we have been describing states of systems by vectors $|\Psi\rangle$ and $\langle\Psi|$, or, more explicitly, by amplitude components $\langle i|\Psi\rangle = \langle\Psi|i\rangle^*$ in some basis:

$$|\Psi\rangle = \sum_i |i\rangle\langle i|\Psi\rangle, \tag{1.3.5a}$$

$$\langle\Psi| = \sum_j \langle\Psi|j\rangle\langle j|. \tag{1.3.5b}$$

However, it will sometimes be very convenient to describe the state of a system instead by an operator or explicitly by a matrix. In such a formalism, the state in Eq. (1.3.5) would be denoted by a single DENSITY OPERATOR $|\Psi\rangle\langle\Psi|$ or DENSITY MATRIX ρ, Eq. (1.3.6), that represents the density operator.

$$|\Psi\rangle\langle\Psi| \leftrightarrow \rho = \begin{pmatrix} \langle 1|\Psi\rangle\langle\Psi|1\rangle & \langle 1|\Psi\rangle\langle\Psi|2\rangle & \cdots & \langle 1|\Psi\rangle\langle\Psi|n\rangle \\ \langle 2|\Psi\rangle\langle\Psi|1\rangle & \langle 2|\Psi\rangle\langle\Psi|2\rangle & \cdots & \langle 2|\Psi\rangle\langle\Psi|n\rangle \\ \vdots & & & \\ \langle n|\Psi\rangle\langle\Psi|1\rangle & \langle n|\Psi\rangle\langle\Psi|2\rangle & \cdots & \langle n|\Psi\rangle\langle\Psi|n\rangle \end{pmatrix}. \tag{1.3.6}$$

We note immediately that ρ is HERMITIAN $\rho_{ij} = \rho_{ji}^*$.

The diagonal terms $\langle i|\Psi\rangle\langle\Psi|i\rangle$ give the quantum mechanical probability that state will be found in base state i. As shown in later chapters this formalism is very useful when one wishes to consider the probabilities associated with thermal averages along with the unavoidable probabilities of quantum mechanics.

1.4 SOME MATRICES USED IN CLASSICAL MECHANICS

The concepts associated with classical matrix applications are generally more down-to-earth than those of the preceding quantum mechanical applications. Nevertheless, Dirac's bra-ket notation is useful in practically any application of matrices. We review now some ways that matrices and Dirac notation will enter our forthcoming discussions of symmetry analysis for classical problems.

A. Force, Mass, and Acceleration Matrices

A classical system of jiggling springs and masses can be described by some number of coordinates $\chi_1, \chi_2, \cdots, \chi_n(t)$ and the same number of velocities $\dot{\chi}_1, \dot{\chi}_2, \cdots, \dot{\chi}_n$. Each mass may need one or more coordinates depending on the nature of the system and its constraints.

Furthermore, we shall suppose that for each coordinate χ_j, there is a corresponding component of applied force F_j on the mass m located at χ_j, and this F_j obeys Newton's equation

$$F_j = m\alpha_j = m\ddot{\chi}_j. \tag{1.4.1}$$

F_j is the component along χ_j of the vector sum of all force applied to m by attached springs.

For example, suppose the two-mass system in Figure 1.4.1 is described by two coordinates (χ_1, χ_2). These coordinates give the positions of mass m_1 and m_2 with respect to their equilibrium positions indicated by dotted lines. Suppose also that the tension of each spring is proportional to the difference between its actual length and its length at equilibrium. (This is Hooke's law.) The constants of proportionality k_j are written over each spring in Figure 1.4.1.

Then the forces F_1 and F_2 are given by the following matrix equation, which relates them to the coordinates χ_1 and χ_2 (note sign!):

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = -\begin{pmatrix} k_1 + k_{12} & -k_{12} \\ -k_{12} & k_2 + k_{12} \end{pmatrix}\begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \equiv -\begin{pmatrix} \mathscr{F}_{11} & \mathscr{F}_{12} \\ \mathscr{F}_{21} & \mathscr{F}_{22} \end{pmatrix}\begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \tag{1.4.2}$$

The matrix in such an equation is called a FORCE MATRIX. This matrix is always Hermitian for a "conservative" system, as we shall prove below.

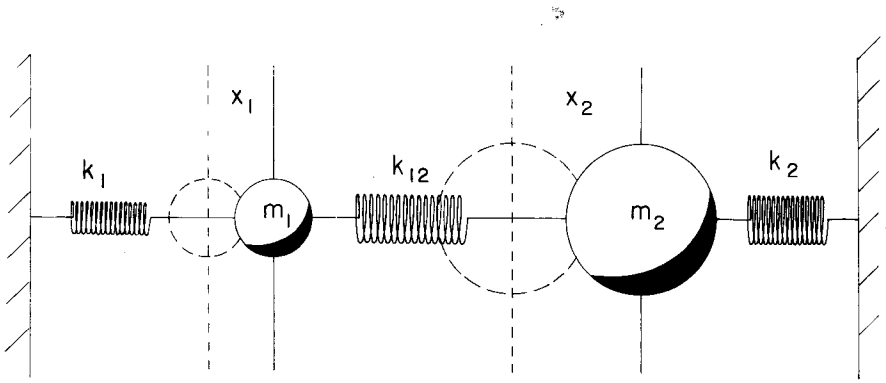


Figure 1.4.1 Example of coupled oscillator system.

Now the forces on each mass are related to acceleration by Newton's equation, Eq. (1.4.1). This equation can be stated in matrix form also, if one defines a mass or INERTIA MATRIX m as in Eq. (1.4.3). This matrix is Hermitian, too, since it is diagonal and real:

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} \equiv \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} \ddot{\chi}_1 \\ \ddot{\chi}_2 \end{pmatrix}. \tag{1.4.3}$$

Sometimes it will be convenient to define a third type of matrix relation involving an ACCELERATION MATRIX $a = -m^{-1}\mathscr{F}$ such as is written in Eq. (1.4.4):

$$\begin{aligned} \begin{pmatrix} \ddot{\chi}_1 \\ \ddot{\chi}_2 \end{pmatrix} &= -\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}^{-1} \begin{pmatrix} \mathscr{F}_{11} & \mathscr{F}_{12} \\ \mathscr{F}_{21} & \mathscr{F}_{22} \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \\ &= -\begin{pmatrix} \dfrac{k_1 + k_{12}}{m_1} & \dfrac{-k_{12}}{m_1} \\ \dfrac{-k_{12}}{m_2} & \dfrac{k_2 + k_{12}}{m_2} \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \\ &= -\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \end{aligned} \tag{1.4.4}$$

It is the eigenvalues and eigenvectors of this matrix that give the ELEMENTARY RESONANT FREQUENCIES and RESONANT MODES of a vibrating system. We can see this very clearly if we reformulate the problem in Dirac notation.

Let us denote the coordinates by $\chi_1 \equiv \langle 1|x\rangle$ and $\chi_2 \equiv \langle 2|x\rangle$ where the "state" $|i\rangle$ stands for that state in which mass i is displaced to $\chi_1 = 1$ unit and the others are at equilibrium $\chi_j = 0$ $(j \neq i)$. State $|x\rangle$ stands for that state in which both masses are off equilibrium by $\langle 1|x\rangle = \chi_1$ and $\langle 2|x\rangle = \chi_2$, respectively. Normalization ($\langle i|i\rangle = 1$) establishes the unit distance.

The mathematical procedures are practically the same as they are for the quantum mechanical Schrodinger equation. [Recall Eqs. (1.1.17)–(1.2.3).] The matrices $\mathscr{F}$, m, and a can all be thought of as representations of the abstract operators **F**, **m**, and **a** in the basis $|1\rangle$, $|2\rangle$. For example, Eq. (1.4.4)

is rewritten as follows:

$$\begin{pmatrix} \langle 1|\ddot{\chi}\rangle \\ \langle 2|\ddot{\chi}\rangle \end{pmatrix} = -\begin{pmatrix} \langle 1|\mathbf{a}|1\rangle & \langle 1|\mathbf{a}|2\rangle \\ \langle 2|\mathbf{a}|1\rangle & \langle 2|\mathbf{a}|2\rangle \end{pmatrix}\begin{pmatrix} \langle 1|x\rangle \\ \langle 2|x\rangle \end{pmatrix}. \tag{1.4.5}$$

Then an eigenbasis $|a_1\rangle, |a_2\rangle$ is found such that the matrix a in this new system is diagonal, with eigenvalues a_1 and a_2:

$$\langle a_i|\mathbf{a}|a_j\rangle = a_i\delta_{ij}, \tag{1.4.6}$$

$$\langle a_i|a_j\rangle = \delta_{ij}, \tag{1.4.7a}$$

$$\sum_i |a_i\rangle\langle a_i| = 1. \tag{1.4.7b}$$

[These vectors can be made according to the techniques of Section 1.2 so they satisfy, along with their left-handed companions $\langle a_1|, \langle a_2|$, Eqs. (1.4.6) and (1.4.7).]

In this basis Eq. (1.4.5) will be very easy to solve since it is uncoupled now:

$$\langle a_j|\ddot{x}\rangle = -a_j\langle a_j|x\rangle.$$

The new coordinates $\langle a_j|x\rangle$, which are often called NORMAL COORDINATES, each oscillate with their particular elementary resonance frequency $\omega_j = \sqrt{a_j}$. There are quite a number of examples of these coordinates in the next few chapters.

For more complicated problems it is sometimes convenient to avoid dealing with the a matrix until the very end of a calculation. One of the problems with a is that, unlike $\mathscr{F}$ and m, it can be non-Hermitian. [See Eq. (1.4.4).] When $a \neq a^\dagger$ then left ($|\ \rangle$) and right ($\langle\ |$) eigenvectors are not related by (†) conjugation (i.e., $\langle a_j| \neq |a_j\rangle^\dagger$) as in Figure 1.2.1.

However, there are other reasons why dealing with an equation like Eq. (1.4.8b) can be easier than working with $\mathbf{a}$ in Eq. (1.4.8a):

$$\mathbf{a}|e_j\rangle = \mathbf{m}^{-1}\mathbf{F}|e_j\rangle = a_j|e_j\rangle, \tag{1.4.8a}$$

$$\mathbf{F}|e_j\rangle = a_j\mathbf{m}|e_j\rangle, \tag{1.4.8b}$$

$$(\mathbf{F} - a_j\mathbf{m})|e_j\rangle = 0. \tag{1.4.8c}$$

This happens when, for various reasons, we want to use a nonorthogonal basis $|A_1\rangle, |A_2\rangle, \ldots$ in which $\langle A_i|A_j\rangle \neq \delta_{ij}$, but $\langle A_i| = |A_i\rangle^\dagger$. In this case we would be seeking an eigenvector $|e_j\rangle$ of the form

$$|e_j\rangle = \varepsilon_1|A_1\rangle + \varepsilon_2|A_2\rangle + \cdots, \tag{1.4.9a}$$

which satisfies a GENERALIZED EIGENVALUE equation

$$\mathbf{F}(\varepsilon_1|A_i\rangle + \varepsilon_2|A_2\rangle + \cdots) = a_j\mathbf{m}(\varepsilon_1|A_1\rangle + \varepsilon_2|A_2\rangle + \cdots). \quad (1.4.9b)$$

Forming the scalar product of this in turn with $\langle A_1| = |A_1\rangle^\dagger$, $\langle A_2| = |A_2\rangle^\dagger$ $\cdots$ results in the following generalized matrix eigenvalue equation in which the force and mass matrices are Hermitian:

$$\begin{pmatrix} \langle A_1|\mathbf{F}|A_1\rangle & \langle A_1|\mathbf{F}|A_2\rangle & \cdots \\ \langle A_2|\mathbf{F}|A_1\rangle & \langle A_2|\mathbf{F}|A_2\rangle & \cdots \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \end{pmatrix}$$

$$= a_j \begin{pmatrix} \langle A_1|\mathbf{m}|A_1\rangle & \langle A_1|\mathbf{m}|A_2\rangle & \cdots \\ \langle A_2|\mathbf{m}|A_1\rangle & \langle A_2|\mathbf{m}|A_2\rangle & \cdots \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \end{pmatrix}. \quad (1.4.10)$$

Now it is easy to see that the left and right eigenvectors $\langle e_j|$ and $|e_j\rangle$ satisfying Eqs. (1.4.11a) and (1.4.11b) must be related by Eq. (1.4.11c):

$$(\mathbf{F} - a_j\mathbf{m})|e_j\rangle = 0, \quad (1.4.11a)$$

$$\langle e_j|(\mathbf{F} - a_j\mathbf{m}) = 0, \quad (1.4.11b)$$

$$\langle e_j| = |e_j\rangle^\dagger. \quad (1.4.11c)$$

The components $(\varepsilon_1\varepsilon_2 \cdots)$ or $(\varepsilon_1^*\varepsilon_2^* \cdots)$ of $|e_j\rangle$ or $\langle e_j|$ are, respectively, proportional to the columns or rows of the adjunct matrix $\langle \mathbf{F} - a_m\mathbf{m}\rangle^{\text{ADJ}}$. [Recall Section 1.2.B(e).]

Rewriting Eq. (1.4.11b) shows the relation between the vector $\langle e_j| = |e_j\rangle^\dagger$ and $\langle a_j|$ in Eq. (1.4.6):

$$\langle e_j|\mathbf{F} = \langle e_j|\mathbf{ma} = a_j\langle e_j|\mathbf{m}. \quad (1.4.12)$$

Equation (1.4.12) indicates that $\langle a_j|$ is proportional to $\langle e_j|\mathbf{m}$, while a similar comparison with Eq. (1.4.11a) shows that $|e_j\rangle$ is proportional to $|a_j\rangle$. We shall choose our normalization of $|e_j\rangle$ such that

$$\langle e_j|\mathbf{m}|e_k\rangle = \delta_{jk}, \quad (1.4.13a)$$

whence

$$\langle e_j|\mathbf{F}|e_k\rangle = a_j\langle e_j|\mathbf{m}|e_k\rangle = a_j\delta_{jk}. \quad (1.4.13b)$$

This $\mathbf{m}$ normalization gives in turn a peculiar form for the completeness

relation:

$$\mathbf{1} = \sum_i |e_i\rangle\langle e_i|\mathbf{m} = \mathbf{1}^\dagger = \sum_i \mathbf{m}|e_i\rangle\langle e_i|. \tag{1.4.14}$$

B. Potential Energy and Hamiltonian Functions

Returning to Eq. (1.4.2), we can see that the work done while changing coordinate χ_1 to $\chi_1 + d\chi_1$ and $\chi_2 + d\chi_2$ is given by Eq. (1.4.15):

$$dW = \begin{pmatrix} -d\chi_1 & d\chi_2 \end{pmatrix}\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \begin{pmatrix} d\chi_1 & d\chi_2 \end{pmatrix}\begin{pmatrix} \mathscr{F}_{11} & \mathscr{F}_{12} \\ \mathscr{F}_{21} & \mathscr{F}_{22} \end{pmatrix}\begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \tag{1.4.15}$$

If, as we supposed, the components $\mathscr{F}_{ij}$ are constants, then we have by integration the following expression for potential energy:

$$\begin{aligned} W \equiv V(\chi_1\chi_2) &= \tfrac{1}{2}\begin{pmatrix} \chi_1 & \chi_2 \end{pmatrix}\begin{pmatrix} \mathscr{F}_{11} & \mathscr{F}_{12} \\ \mathscr{F}_{21} & \mathscr{F}_{22} \end{pmatrix}\begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \\ &= \tfrac{1}{2}\sum_i\sum_j \langle x|i\rangle\langle i|\mathbf{F}|j\rangle\langle j|x\rangle = \tfrac{1}{2}\langle x|\mathbf{F}|x\rangle. \end{aligned} \tag{1.4.16}$$

Note that a Taylor expansion of a general potential function $V(\chi_1\chi_2)$ gives a much more complicated result. (We ignore the zeroth- and first-order terms which can be taken to be zero if $\chi_i = 0$ is the equilibrium point.)

$$\begin{aligned} V(\chi_1\chi_2) = {} & \frac{1}{2!}\sum_{ij}\chi_i\chi_j\frac{\partial^2 V}{\partial\chi_i\,\partial\chi_j}\bigg|_{\chi=0} \\ & + \frac{1}{3!}\sum_{ijk}\chi_i\chi_j\chi_k\frac{\partial^3 V}{\partial\chi_i\,\partial\chi_j\,\partial\chi_k}\bigg|_{\chi=0} + \cdots. \end{aligned} \tag{1.4.17}$$

However, if only the quadratic terms are nonzero, then we prove that F is Hermitian for potential-driven or conservative systems:

$$\mathscr{F}_{ij} = \frac{\partial^2 V}{\partial\chi_i\,\partial\chi_j}\bigg|_{\chi=0} = \mathscr{F}_{ji}. \tag{1.4.18}$$

A function $H(q, p)$, called the HAMILTONIAN FUNCTION, can be constructed in terms of canonical momenta $p(j)$ and coordinates $q(j)$ as defined in the following. Equations (1.4.13) and (1.4.14) are used.

$$\begin{aligned} H(q,p) &= T + V = \tfrac{1}{2}\langle\dot{x}|\mathbf{m}|\dot{x}\rangle + \tfrac{1}{2}\langle x|\mathbf{F}|x\rangle \\ &= \tfrac{1}{2}\sum_j\left[\langle\dot{x}|\mathbf{m}|e_j\rangle\langle e_j|\mathbf{m}|\dot{x}\rangle + \langle x|\mathbf{m}|e_j\rangle\langle e_j|\mathbf{F}|x\rangle\right]. \end{aligned}$$

This may be written as

$$H(q,p) = \tfrac{1}{2}\sum_j \left[p(j)^2 + a_j |q(j)|^2 \right], \tag{1.4.19a}$$

where

$$q(j) = \langle e_j | \mathbf{m} | x \rangle = \langle a_j | x \rangle, \tag{1.4.19b}$$

and

$$p(j) = \langle e_j | \mathbf{m} | \dot{x} \rangle = \dot{q}(j). \tag{1.4.19c}$$

It is easy to see that the latter obey Hamilton's equations:

$$\frac{\partial H}{\partial q(j)} = a_j q(j) = -\dot{p}(j), \qquad \frac{\partial H}{\partial p(j)} = p(j) = \dot{q}(j). \tag{1.4.20}$$

This is a necessary prerequisite for any quantum theory in which each momentum $p(j)$ is replaced by an operator $(\hbar/i)\partial/\partial q(j)$ to make the Hamiltonian operator into the Schrödinger equation.

However, for classical vibration problems it is often more convenient to deal directly with Newton's equations. The F matrix is usually easy to obtain by inspection. For example, for the system in Figure 1.4.2 we would obtain the component $\langle i|\mathbf{F}|j\rangle$ by simply computing the product of the projections of each spring on coordinate axis $\chi_i = \langle i|x\rangle$ and $\chi_j = \langle j|x\rangle$ for all springs they share, and summing these products:

$$\begin{aligned} -F_1 &= \langle 1|\mathbf{F}|1\rangle\langle 1|x\rangle + \langle 1|\mathbf{F}|2\rangle\langle 2|x\rangle \\ &\quad + \langle 1|\mathbf{F}|3\rangle\langle 3|x\rangle + \langle 1|\mathbf{F}|4\rangle\langle 4|x\rangle + \cdots, \\ F_1 &= \left(k_1 \cos^2\phi_1 + k_1' \cos^2\phi_1' + k_{12}\cos^2\theta_1\right)\chi_1 \\ &\quad + \left(k_1 \cos\phi_2 \cos\phi_1 + k_2' \cos\phi_1 + k_{12}\cos\theta_2\cos\theta_1\right)\chi_2 \\ &\quad + \left(k_{12}\cos\theta_3\cos\theta_1\right)\chi_3 + \left(k_{12}\cos\theta_4\cos\theta_1\right)\chi_4 + \cdots. \end{aligned}$$

One should always keep in mind that classical equations such as Eqs. (1.4.2) and (1.4.4) are approximate descriptions which assume small χ_j. For large χ_j one will need to include anharmonic $\chi_j^3, \chi_j\chi_k^2, \chi_j^4 \ldots$, etc., to model the behavior of a highly stretched spring and changing geometry. This is especially true for more complicated problems in two or three dimensions, as is represented by Figure 1.4.2, or for models of molecular systems.

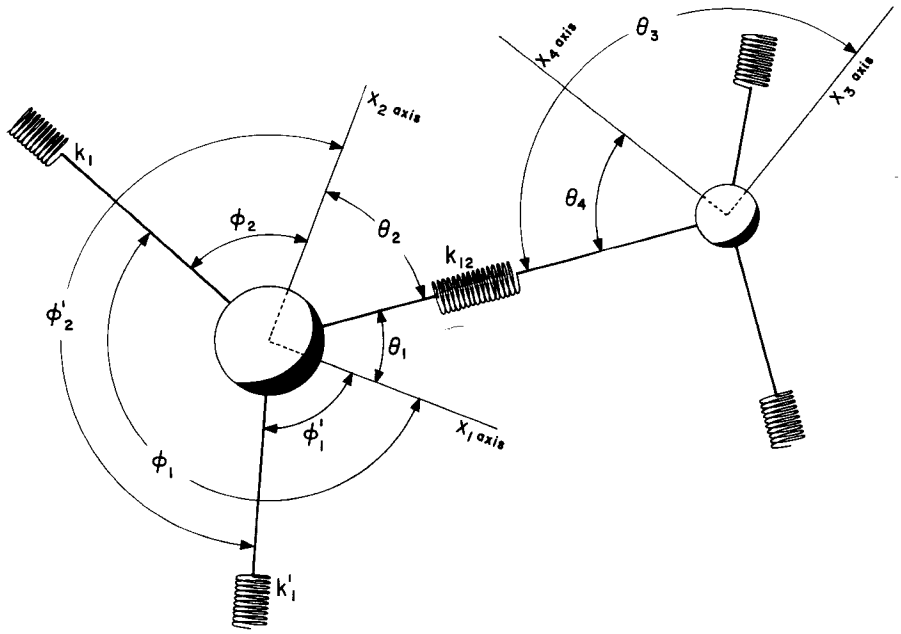


Figure 1.4.2 Example of more complicated spring-mass system.

1.5 WAVE FUNCTIONAL ANALYSIS AND CONTINUOUS VECTOR SPACES

Dirac notation was designed to be used in vector spaces that had a continuous infinity of dimensions. The symbol $\langle x|\Psi\rangle$ is Dirac's notation for a complex wavefunction $\Psi(x)$ of a continuous coordinate x. The complex conjugate $\Psi^*(x)$ is the reversed bra-ket $\langle\Psi|x\rangle$:

$$\langle x|\Psi\rangle = \Psi(x) \qquad \langle\Psi|x\rangle = \Psi^*(x). \tag{1.5.1}$$

Several modifications of the formalism are needed to accommodate this continuously infinite ket vector space $\{\cdots|x\rangle\cdots\}$ or its conjugate bra vector space $\{\cdots\langle x|\cdots\}$. These are known in mathematics as *Banach spaces* and the description of wavefunctions by these vectors falls under the mathematical heading of *functional analysis*. This is a large and complicated subject. Nevertheless, the basic ideas of completeness and orthogonality remain intact. A comparison between discrete and continuous vector spaces is given now.

A. Functional Scalar Products

Scalar products between continuously infinite bra and ket vectors involve what is known as Dirac's delta function $\delta(x, y)$. In place of the usual Kronecker delta δ_{ab} relation,

$$\langle a|b\rangle = \delta_{ab} = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b, \end{cases} \tag{1.5.2a}$$

one uses the Dirac delta relation,

$$\langle x|y\rangle = \delta(x, y) = \begin{cases} \infty & \text{if } x = y \\ 0 & \text{if } x \neq y. \end{cases} \tag{1.5.2b}$$

Replacing unity (1) by infinity (∞) in the scalar product is necessary because the sum over discrete dimensions is replaced by an integral. It is the sum or integral that needs to equal unity. It is easy to see that

$$\sum_b \langle a|b\rangle = \sum_b \delta_{ab} = 1,$$

but the corresponding Dirac sum is not so obvious:

$$\int dy \langle x|y\rangle = \int dy\, \delta(x, y) = 1.$$

The delta function is zero everywhere except at $x = y$ where it goes to infinity in such a way as to have unit area. Both kinds of delta functions are designed to extract a particular component from a sum. The following discrete sum

$$\langle a|\Psi\rangle = \sum_b \langle a|b\rangle\langle b|\Psi\rangle = \sum_b \delta_{ab}\Psi_b = \Psi_a \tag{1.5.3a}$$

is analogous to the following integral:

$$\langle x|\Psi\rangle = \int dy \langle x|y\rangle\langle y|\Psi\rangle = \int dy\, \delta(x, y)\Psi(y) = \Psi(x). \tag{1.5.3b}$$

B. Orthonormality and Completeness

The completeness relation (1.2.19a) for a discrete space $\{|1\rangle, |2\rangle \cdots\}$ has the form

$$\mathbf{1} = \sum_b |b\rangle\langle b|. \tag{1.5.4a}$$

It is replaced by the following integral for a continuous space $\{ \cdots |y\rangle \cdots \}$:

$$\mathbf{1} = \int dy |y\rangle\langle y| . \tag{1.5.4b}$$

For many applications to quantum mechanics there is a discrete set of eigenstates $\{|\Psi_1\rangle, |\Psi_2\rangle \cdots\}$ but their wavefunctions range over a continuous set of position states $\{ \cdots |x\rangle \cdots \}$. The wavefunctions $\Psi_a(x)$ can be thought of as components of a transformation matrix $\langle x|\Psi_a\rangle$ between these two bases. Orthogonality and completeness of the transformation is analogous to the discrete space relations (1.2.19b) and (1.2.20), respectively,

$$\langle x|y\rangle = \delta(x, y) = \sum_a \langle x|\Psi_a\rangle\langle\Psi_a|y\rangle,$$

$$\langle \Psi_a|\Psi_b\rangle = \delta_{ab} = \int dz \langle\Psi_a|z\rangle\langle z|\Psi_b\rangle. \tag{1.5.5a}$$

Here is the same equation in wavefunction notation:

$$\langle x|y\rangle = \delta(x, y) = \sum_a \Psi_a^*(x)\Psi_a(y),$$

$$\langle \Psi_a|\Psi_b\rangle = \delta_{ab} = \int dz\, \Psi_a^*(z)\Psi_b(z). \tag{1.5.5b}$$

Other applications to quantum mechanics involve a continuous set of eigenstates $\{ \cdots |\Psi_k\rangle \cdots \}$ defined by wavefunctions over a continuous set of position states $\{ \cdots |x\rangle \cdots \}$. For example, plane wave functions $\Psi_k(x) = e^{ikx}/\sqrt{2\pi}$ can be thought of as components of a Fourier transformation matrix $\langle x|\Psi_k\rangle$ between position and momentum space. The orthogonality and completeness of this transformation are similar relations:

$$\langle x|y\rangle = \delta(x, y) = \int dk \langle x|\Psi_k\rangle\langle\Psi_k|y\rangle,$$

$$\langle \Psi_k|\Psi_{k'}\rangle = \delta(k, k') = \int dz \langle\Psi_k|z\rangle\langle z|\Psi_{k'}\rangle. \tag{1.5.6a}$$

Here is the same equation for plane waves:

$$\langle x|y\rangle = \delta(x, y) = \frac{1}{2\pi}\int dk\, e^{-ikx}e^{iky},$$

$$\langle \Psi_k|\Psi_{k'}\rangle = \delta(k, k') = \frac{1}{2\pi}\int dz\, e^{-ikz}e^{ik'z}. \tag{1.5.6b}$$

Orthonormality for one basis is the same type of equation as completeness for the other.

C. Differential Operators and Conjugates

Analogies between discrete and continuous vector spaces can be extended to include linear differential operators. For example, the following differential operator M acting on a function $f(x)$,

$$Mf(x) = \mu(x)\frac{d^2f}{dx^2} + \nu(x)\frac{df}{dx} + \lambda(x)f \tag{1.5.7}$$

can be expressed as a Dirac matrix element

$$\langle x|\mathbf{M}|f\rangle = \int dy \langle x|\mathbf{M}|y\rangle\langle y|f\rangle, \tag{1.5.8}$$

where $\langle x|\mathbf{M}|y\rangle$ is expressed in terms of Dirac delta function derivatives.

$$\langle x|\mathbf{M}|y\rangle = \mu(x)\frac{d^2\delta(x,y)}{dy^2} + \nu(x)\frac{d\delta(x,y)}{dy} + \lambda(x)\delta(x,y). \tag{1.5.9}$$

The Dirac delta derivative extracts the value of a derivative of a function, that is, the value of the function minus its value at an infinitesimally nearby point. Integration by parts gives

$$\int_{-\infty}^{\infty} dx\, f(x)\frac{d\delta(x,y)}{dx} = f(x)\delta(x,y)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx\, \frac{df(x)}{dx}\delta(x,y) = -f'(y),$$

$$\int_{-\infty}^{\infty} dy\, f(y)\frac{d\delta(x,y)}{dy} = f'(x). \tag{1.5.10}$$

The transpose conjugate operator $\mathbf{M}^\dagger$ has matrix element $\langle x|\mathbf{M}^\dagger|y\rangle = \langle y|\mathbf{M}|x\rangle^*$:

$$\begin{aligned}\langle x|\mathbf{M}^\dagger|y\rangle &= \langle y|\mathbf{M}|x\rangle^* \\ &= \mu^*(y)\frac{d^2\delta(y,x)}{dy^2} + \nu^*(y)\frac{d\delta(y,x)}{dy} + \lambda^*(y)\delta(y,x).\end{aligned} \tag{1.5.11}$$

The result of applying the operator $\mathbf{M}^\dagger$ using (1.5.8) and (1.5.11) is the following,

$$\mathbf{M}^\dagger f(x) = \frac{d^2}{dx^2}(\mu^*(x)f(x)) - \frac{d}{dx}(\nu^*(x)f(x)) + \lambda^*(x)f(x). \tag{1.5.12}$$

This is called the *adjoint* differential operator. If the operator is self-adjoint ($\mathbf{M} = \mathbf{M}^\dagger$) as are all Hamiltonian operators then the operation above is equal

to the original one. This places restrictions on the form of the functions $\mu(x)$, $\nu(x)$, and $\lambda(x)$.

D. Resolvants

Given an operator $\mathbf{M}$ there is a useful function of that operator which is called the *resolvant* $R_\lambda(\mathbf{M}) = -Q_\lambda(\mathbf{M})$. It uses a parameter λ,

$$Q_\lambda(\mathbf{M}) = \frac{1}{\lambda\mathbf{1} - \mathbf{M}} = -R_\lambda(\mathbf{M}). \tag{1.5.13}$$

If $\lambda = 0$, the first resolvant R_0 is the inverse $\mathbf{M}^{-1}$ of $\mathbf{M}$ and Q_0 is the negative inverse $-\mathbf{M}^{-1}$,

$$R_0(\mathbf{M}) = \mathbf{M}^{-1}.$$

The resolvant $Q_\lambda(\mathbf{M})$ expands into a series of $\mathbf{M}/\lambda$ powers,

$$Q_\lambda(\mathbf{M}) = \frac{1}{\lambda(\mathbf{1} - \mathbf{M}/\lambda)} = \frac{1}{\lambda}\left(\mathbf{1} + \frac{\mathbf{M}}{\lambda} + \frac{\mathbf{M}^2}{\lambda^2} + \cdots\right).$$

Applying Cauchy's integral theorem to the complex variable λ yields contour integral expressions for each term,

$$\frac{1}{2\pi i}\oint_L d\lambda \frac{1}{\lambda\mathbf{1} - \mathbf{M}} = \mathbf{1}, \qquad \frac{1}{2\pi i}\oint_L d\lambda \frac{\lambda}{\lambda\mathbf{1} - \mathbf{M}} = \mathbf{M},$$
$$\frac{1}{2\pi i}\oint_L d\lambda \frac{\lambda^2}{\lambda\mathbf{1} - \mathbf{M}} = \mathbf{M}^2, \ldots. \tag{1.5.14}$$

This gives a symbolic expression for any function $f(\mathbf{M})$ of the operator $\mathbf{M}$ in terms of an integral of the function $f(\lambda)$ times the resolvant:

$$f(\mathbf{M}) = \frac{1}{2\pi i}\oint_L d\lambda \frac{f(\lambda)}{\lambda\mathbf{1} - \mathbf{M}} = \frac{1}{2\pi i}\oint_L d\lambda\, f(\lambda) Q_\lambda(\mathbf{M}). \tag{1.5.15}$$

To use this we need to know where the poles of the resolvant $Q_\lambda(\mathbf{M})$ are. It is a rational function of $\lambda\mathbf{1} - \mathbf{M}$ divided by the secular polynomial $S(\lambda) = \det|\lambda\mathbf{1} - \mathbf{M}|$,

$$Q_\lambda(\mathbf{M}) = (\lambda\mathbf{1} - \mathbf{M})^{-1} = \frac{(\lambda\mathbf{1} - \mathbf{M})^{\mathrm{ADJ}}}{\det|\lambda\mathbf{1} - \mathbf{M}|} = \frac{a(\lambda, \mathbf{M})}{S(\lambda)}. \tag{1.5.16}$$

This implies that the resolvant's poles are located at its eigenvalues

$\{\lambda_1, \lambda_2, \lambda_3, \ldots\}$, that is, the roots of $S(\lambda) = 0$. The adjunct function can have no singularities of its own. This means the contour loop L in the contour integral equations above can be deformed into a set of small circles $\{l_1, l_2, \ldots\}$ surrounding each distinct eigenvalue in the complex plane. Each integral then breaks into a sum over distinct eigenvalues.

The first integral expression for the unit operator becomes an operator completeness relation like (1.2.18) which we discussed before:

$$\mathbf{1} = \frac{1}{2\pi i}\oint_L d\lambda \, \frac{1}{\lambda \mathbf{1} - \mathbf{M}} = \sum_{\lambda_a} \mathbf{P}_{\lambda_a}. \tag{1.5.17}$$

Here the projection operators have the form

$$\mathbf{P}_{\lambda_a} = \frac{1}{2\pi i}\oint_{l_a} d\lambda \, \frac{1}{\lambda \mathbf{1} - \mathbf{M}} = \frac{1}{2\pi i}\oint_{l_a} d\lambda \, Q_\lambda(\mathbf{M}). \tag{1.5.18}$$

If all the eigenvalues are distinct then each eigenvalue is a simple pole. In that case it can be shown that the **P**-operators are orthogonal idempotents as in Eq. (1.2.17). Then the **M** operator and its powers have spectral decompositions like Eq. (1.2.21):

$$\mathbf{M} = \frac{1}{2\pi i}\int_L d\lambda \, \frac{\lambda}{\lambda \mathbf{1} - \mathbf{M}} = \sum_{\lambda_a} \lambda_a \mathbf{P}_{\lambda_a}. \tag{1.5.19}$$

However, for degenerate eigenvalue ($\lambda_a = \lambda_b$) it may be necessary to include extra terms in a spectral decomposition such as in the following:

$$\begin{aligned} \mathbf{M} &= \sum_{\lambda_a} \lambda_a \frac{1}{2\pi i}\oint_{l_a} d\lambda \, \frac{\lambda}{\lambda \mathbf{1} - \mathbf{M}} \\ &= \sum_{\lambda_a} \left[\lambda_a \frac{1}{2\pi i}\oint_{l_a} d\lambda \, \frac{1}{\lambda \mathbf{1} - \mathbf{M}} + \frac{1}{2\pi i}\oint_{l_a} d\lambda \, \frac{\lambda - \lambda_a}{\lambda \mathbf{1} - \mathbf{M}} \right] \\ &= \sum_{\lambda_a} (\mathbf{P}_{\lambda_a} + \mathbf{N}_{\lambda_a}), \quad \text{where: } \mathbf{N}_{\lambda_a} = \frac{1}{2\pi i}\oint_{l_a} d\lambda \, \frac{\lambda - \lambda_a}{\lambda \mathbf{1} - \mathbf{M}}. \end{aligned} \tag{1.5.20}$$

The extra term $\mathbf{N}_\lambda$ is zero if the resolvant still has a simple pole at that eigenvalue. Otherwise it is a nilpotent operator, that is, some power of it is zero. If the eigenvalue is a second order pole of the resolvant, then $(\mathbf{N}_\lambda)^2$ is zero:

$$(\mathbf{N}_{\lambda_a})^2 = \frac{1}{2\pi i}\oint_{l_a} d\lambda \, \frac{(\lambda - \lambda_a)^2}{\lambda \mathbf{1} - \mathbf{M}} = 0. \tag{1.5.21}$$

Resolvants provide a powerful tool for spectral analysis of operators.

APPENDIX A. ELEMENTARY VECTOR NOTATIONS AND THEORY

In Chapter 1 we introduced Dirac's vector notation $|\ \rangle$ and $\langle\ |$. Here we relate this notation to the older notation of vector analysis.

The dirac vector in Eq. (A.1a) can be written in the older notation as Eq. (A.1b):

$$|v\rangle = |x\rangle\langle x|v\rangle + |y\rangle\langle y|v\rangle, \tag{A.1a}$$

$$\mathbf{v} = \hat{x}v_x + \hat{y}v_y, \tag{A.1b}$$

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} \langle x|v\rangle \\ \langle y|v\rangle \end{pmatrix}. \tag{A.1c}$$

It can be represented by a column vector in Eq. (A.1c) or as is shown also in Eq. (1.1.11). Figure A.1 gives a geometrical representation of the vectors $\mathbf{v}$ and $\mathbf{w}$ in a two-dimensional space where their properties can be seen clearly.

In Figure A.2 we show the meaning of multiplying a vector by a scalar by comparing the vectors $\mathbf{v}$, $2\mathbf{v}$, and $-\mathbf{v}$, where $\mathbf{v} = -3\hat{x} + 4\hat{y}$.

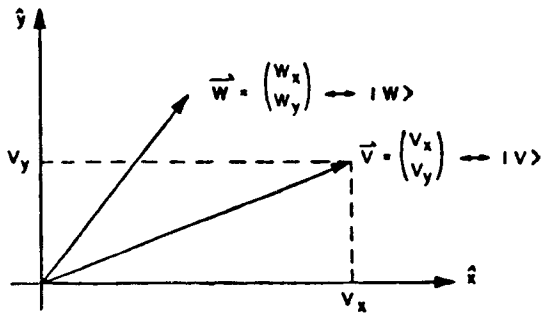


Figure A.1

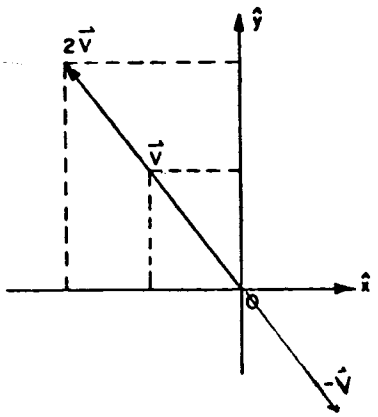


Figure A.2

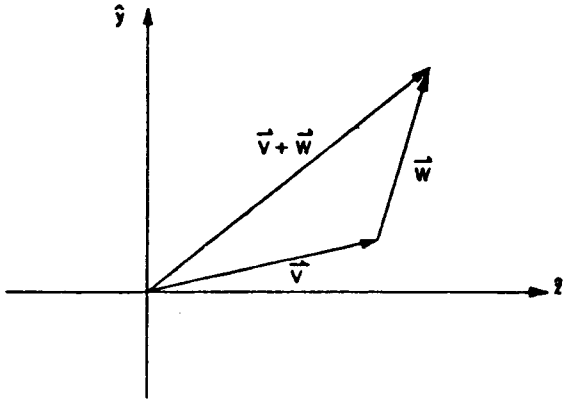


Figure A.3

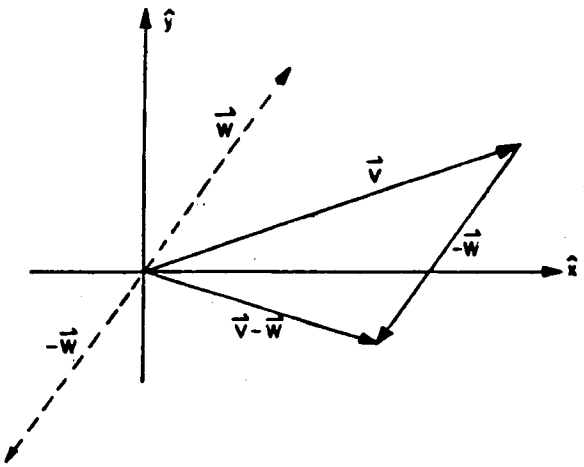


Figure A.4

The geometrical interpretation of the addition of two vectors, Eq. (A.2), is shown in Figure A.3, and Figure A.4 shows the formation of $\mathbf{v} - \mathbf{w}$:

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} v_x \\ v_y \end{pmatrix} + \begin{pmatrix} w_x \\ w_y \end{pmatrix} = \begin{pmatrix} v_x + w_x \\ v_y + w_y \end{pmatrix}. \tag{A.2}$$

Any two vectors in our two-dimensional space that are not proportional to each other could, by linear combination, form all other vectors in the space. This is an important concept that can be generalized to n-dimensional space.

The length or magnitude $|\mathbf{v}|$ of a vector $\mathbf{v}$ is given by the Pythagorian theorem [Eq. (A.3)]:

$$|\mathbf{v}| = \left(\nu_x^2 + \nu_y^2\right)^{1/2}. \tag{A.3}$$

Equation (A.3) is the square root of the scalar product defined by Eq. (1.1.13) and Eq. (A.4):

$$|\mathbf{v}| = (\mathbf{v} \cdot \mathbf{v})^{1/2} = \begin{pmatrix} \nu_x & \nu_y \end{pmatrix} \begin{pmatrix} \nu_x \\ \nu_y \end{pmatrix} = \left(\nu_x^2 + \nu_y^2\right)^{1/2}. \tag{A.4}$$

The geometrical significance of the scalar product of two arbitrary vectors $\mathbf{v}$ and $\mathbf{w}$ in Figure A.5 is given now.

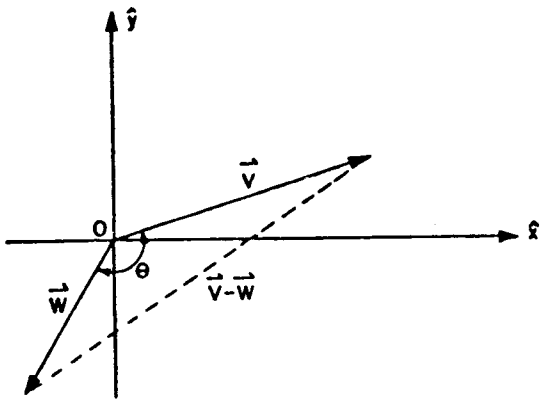


Figure A.5

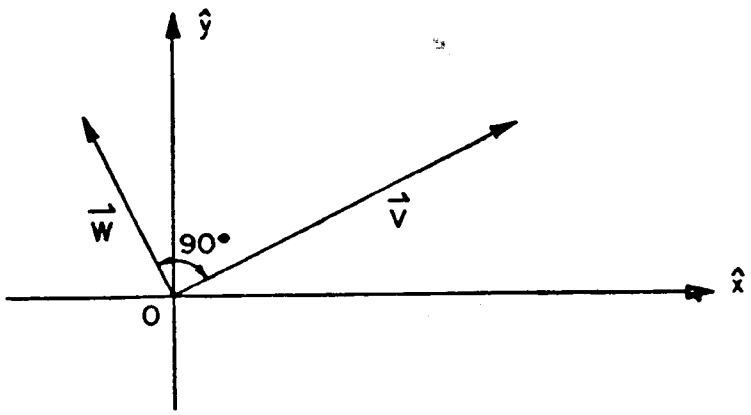


Figure A.6

Applying the law of cosines to the triangle described by the two vectors we arrive at Eq. (A.5a):

$$|\mathbf{v} - \mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 - 2|\mathbf{v}|\,|\mathbf{w}| \cos\theta,$$
$$(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2|\mathbf{v}|\,|\mathbf{w}| \cos\theta. \tag{A.5a}$$

Expanding the left-hand side of Eq. (A.5a), we have

$$(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2\mathbf{v} \cdot \mathbf{w}, \tag{A.5b}$$

and comparing Eq. (A.5a) with Eq. (A.5b), we derive the desired geometrical interpretation of $\mathbf{v} \cdot \mathbf{w}$, which is given in Eq. (A.6):

$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}|\,|\mathbf{w}| \cos\theta. \tag{A.6}$$

To require that the vectors v and w be normalized simply means that they must be unit vectors as in Eq. (A.7):

$$\mathbf{w} \cdot \mathbf{w} = 1 = \mathbf{v} \cdot \mathbf{v}. \tag{A.7}$$

The vectors $\mathbf{x}_1\mathbf{x}_2 \cdots \mathbf{x}_n$ are said to by LINEARLY INDEPENDENT if the relation

$$\sum_{i=1} a_i \mathbf{x}_i = 0 \tag{A.8}$$

necessarily implies that all $a_i = 0$.

If in Eq. (A.8) there exist at least two nonvanishing a_i, then it is said that the vectors $\mathbf{x}_1\mathbf{x}_2 \cdots \mathbf{x}_n$ are linearly dependent.

The maximum number of linearly independent vectors in a space is called the dimension of the space.

For completeness, we present the mathematical definitions of a linear vector space.

Given that $|i\rangle$, $|j\rangle$, and $|k\rangle$ are vectors in the vector space and α and β are complex numbers, then the following is true:

(1) Given any two vectors in the space then their sum is in the space. This is called closure.

(2) The commutative law of addition holds; i.e.,

$$|i\rangle + |j\rangle = |j\rangle + |i\rangle. \tag{A.9a}$$

(3) The associative law of addition holds; i.e.,

$$(|i\rangle + |j\rangle) + |k\rangle = |i\rangle + (|j\rangle + |k\rangle). \tag{A.9b}$$

(4) There exists a null element 0 such that

$$0 + |i\rangle = |i\rangle + 0 = |i\rangle. \tag{A.9c}$$

(5) There exists an element $|i'\rangle$ for every $|i\rangle$ such that

$$|i\rangle + |i'\rangle = 0. \tag{A.9d}$$

(6) The associative law of multiplication holds; i.e.,

$$\alpha(\beta|i\rangle) = \alpha\beta|i\rangle. \tag{A.9e}$$

(7) The distributive law with respect to the addition of complex number holds; i.e.,

$$(\alpha + \beta)|i\rangle = \alpha|i\rangle + \beta|i\rangle. \tag{A.9f}$$

(8) The distributive law with respect to the additions of vectors holds; i.e.,

$$\alpha(|i\rangle + |j\rangle) = \alpha|i\rangle + \alpha|j\rangle. \tag{A.9g}$$

The mathematical definitions of the inner or scalar product follow:

(1) $$\langle i|(|j\rangle + |k\rangle) = \langle i|j\rangle + \langle i|k\rangle. \tag{A.10a}$$

(2) $$\langle i|(\alpha|j\rangle = \alpha\langle i|j\rangle. \tag{A.10b}$$

(3) $$\langle i|j\rangle = \langle j|i\rangle^*. \tag{A.10c}$$

(4) $$\langle i|i\rangle \geq 0. \tag{A.10d}$$

APPENDIX B. LINEAR EQUATIONS, MATRICES, DETERMINANTS, AND INVERSES

The simplest nontrivial example of a matrix equation is the linear equation, Eq. (B.1), involving two unknowns x_1 and x_2 (we assume the quantities $\mathcal{M}_{ij}$ and y_n are known constants). This is written in three ways to display the matrix notation:

$$\begin{aligned} \mathcal{M}_{11}x_1 + \mathcal{M}_{12}x_2 &= y_1, \\ \mathcal{M}_{21}x_1 + \mathcal{M}_{22}x_2 &= y_2, \end{aligned} \tag{B.1a}$$

$$\begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \tag{B.1b}$$

$$\mathcal{M}x = y. \tag{B.1c}$$

This equation is solved if and when the INVERSE $\mathcal{M}^{-1}$ of matrix $\mathcal{M}$ is found such that Eq. (B.2) holds, where 1 is the IDENTITY matrix:

$$\mathcal{M}^{-1}\mathcal{M} = \begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \equiv 1. \tag{B.2}$$

This will give a solution because the product $(1 \cdot x)$ of the identity matrix with any vector x is equal to x, and so Eqs. (B.1) and (B.2) together give Eq. (B.3):

$$\mathcal{M}^{-1}\mathcal{M}x = 1x = x = \mathcal{M}^{-1}y. \tag{B.3}$$

Simple algebra gives, for the inverse of $\mathcal{M}$,

$$\begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \dfrac{\mathcal{M}_{22}}{\mathcal{M}_{11}\mathcal{M}_{22} - \mathcal{M}_{12}\mathcal{M}_{21}} & \dfrac{-\mathcal{M}_{12}}{\mathcal{M}_{11}\mathcal{M}_{22} - \mathcal{M}_{12}\mathcal{M}_{21}} \\ \dfrac{-\mathcal{M}_{21}}{\mathcal{M}_{11}\mathcal{M}_{22} - \mathcal{M}_{22}\mathcal{M}_{21}} & \dfrac{\mathcal{M}_{11}}{\mathcal{M}_{11}\mathcal{M}_{22} - \mathcal{M}_{12}\mathcal{M}_{21}} \end{pmatrix}.$$

We derive now a general formula for the inverse of an $n \times n$ matrix.

Such a formula follows from the properties of the DETERMINANT (det $\mathcal{M}$) of a matrix $\mathcal{M}$. The determinant is defined by Eq. (B.4), where, for convenience, we use a different notation for matrix components:

$$\begin{aligned} \det \mathcal{M} &= \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 & \cdots \\ b_1 & b_2 & b_3 & b_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix} \\ &= \sum_{\substack{\text{permutations } P \\ \text{of } 1\,2\,3\cdots n}} (-1)^{\text{parity}} P(a_1 b_2 c_3 d_4 \cdots) \\ &= (a_1 b_2 c_3 d_4 \cdots) - (a_1 b_3 c_2 d_4 \cdots) + (a_1 \cdots \\ &\quad -(a_2 b_1 c_3 d_4 \cdots) + (a_2 b_3 c_1 d_4 \cdots) + (a_2 \cdots \\ &\quad +(a_3 b_1 c_2 d_4 \cdots) - (a_3 b_2 c_1 d_4 \cdots) + (a_3 \cdots \\ &\quad \vdots \,. \end{aligned} \tag{B.4}$$

A permutation operation P in Eq. (B.4) is said to have parity $= -1$ if it is accomplished by an odd number of "hops" of one number over another

$\cdots ijkl \cdots \rightarrow \cdots ik\overleftarrow{jl} \cdots$ For example, in permuting $4123\cdots$ to $1234\cdots$ we use three hops: $4123 \rightarrow 1\overleftarrow{4}23 \rightarrow 1\overleftarrow{2}43 \rightarrow 12\overleftarrow{3}4$. On the other hand a permutation is said to have parity $= +1$ if it is done in an even number of hops.

From Eq. (B.4) we can understand the expansion of a determinant in terms of smaller subdeterminants called MINORS, as shown in the following. We abbreviate the sum over permutation by $\Sigma(-)$.

$$\det \mathscr{M} = \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 & \cdots \\ b_1 & b_2 & b_3 & b_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix}$$

$$= a_1 \sum_{2,3,4,\ldots,n} (-)(b_2 c_3 d_4 \cdots) - a_2 \sum_{1,3,4,\ldots,n} (-)(b_1 c_3 d_4 \cdots) + a_3 \sum_{1,2,4,\ldots,n} (-)(b_1 c_2 d_4 \cdots) \cdots$$

$$= a_1 \det \begin{vmatrix} b_2 & b_3 & b_4 & \cdots \\ c_2 & c_3 & c_4 & \cdots \\ d_2 & d_3 & d_4 & \cdots \end{vmatrix} - a_2 \det \begin{vmatrix} b_1 & b_3 & b_4 & \cdots \\ c_1 & c_3 & c_4 & \cdots \\ d_1 & d_3 & d_4 & \cdots \end{vmatrix} + a_3 \det \begin{vmatrix} b_1 & b_2 & b_4 & \cdots \\ c_1 & c_2 & c_4 & \cdots \\ d_1 & d_2 & d_4 & \cdots \end{vmatrix}$$

$$= a_1 \mu_{11} - a_2 \mu_{12} + a_3 \mu_{13} - \cdots . \tag{B.5}$$

A minor μ_{ij} of a matrix $\mathscr{M}$ is the determinant obtained from $\mathscr{M}$ after erasing its ith row and its jth column. In Eq. (B.5) we used the minors of the first row of $\mathscr{M}$. However, any row or column of $\mathscr{M}$ could be similarly used to give the general equation (B.6):

$$\det \mathscr{M} = \sum_i (-1)^{i+j} \mathscr{M}_{ij} \mu_{ij}, \qquad j = 1, 2, \ldots, \text{ or } n, \tag{B.6a}$$

$$\det \mathscr{M} = \sum_j (-1)^{i+j} \mathscr{M}_{ij} \mu_{ij}, \qquad i = 1, 2, \ldots, \text{ or } n. \tag{B.6b}$$

Now suppose we define a matrix called the ADJUNCT matrix $\mathscr{M}^{\text{adj}}$ of $\mathscr{M}$ by Eq. (B.7) (note switching of i and j in μ_{ji}!):

$$\mathscr{M}^{\text{adj}}_{ij} = (-1)^{i+j} \mu_{ji}. \tag{B.7}$$

$\mathcal{M}^{\text{adj}}$ together with the original matrix $\mathcal{M}$ obey Eq. (B.8), as can be seen by studying an explicit representation of it in Eq. (B.8c).

$$\sum_j \mathcal{M}_{ij}\mathcal{M}_{jk}^{\text{adj}} = (\det \mathcal{M})\delta_{ik}, \tag{B.8a}$$

$$\mathcal{M}\mathcal{M}^{\text{adj}} = (\det \mathcal{M})1, \tag{B.8b}$$

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & \cdots \\ b_1 & b_2 & b_3 & b_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \end{pmatrix}$$

$$\times \begin{pmatrix} \Sigma(-)b_2c_3d_4 \cdots & -\Sigma(-)a_2c_3d_4 \cdots & \Sigma(-)a_2b_3d_4 \cdots \\ -\Sigma(-)b_1c_3d_4 \cdots & \Sigma(-)a_1c_3d_4 \cdots & -\Sigma(-)a_1b_3d_4 \cdots \\ \Sigma(-)b_1c_2d_4 \cdots & -\Sigma(-)a_1c_2d_4 \cdots & \Sigma(-)a_1b_2d_4 \cdots \\ -\Sigma(-)b_1c_2d_3 \cdots & \Sigma(-)a_1c_2d_3 \cdots & -\Sigma(-)a_1b_2d_3 \cdots \end{pmatrix}$$

$$\begin{pmatrix} \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 & \cdots \\ b_1 & b_2 & b_3 & b_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix} \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 & \cdots \\ a_1 & a_2 & a_3 & a_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix} \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ a_1 & a_2 & a_3 & a_4 \\ d_1 & d_2 & d_3 & d_4 \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix} \cdots \\ \det \begin{vmatrix} b_1 & b_2 & b_3 & b_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix} \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 & \cdots \\ b_1 & b_2 & b_3 & b_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix} \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ b_1 & b_2 & b_3 & b_4 \\ d_1 & d_2 & d_3 & d_4 \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix} \cdots \\ \det \begin{vmatrix} c_1 & c_2 & c_4 & c_4 & \cdots \\ b_1 & b_2 & b_3 & b_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix} \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ d_1 & d_2 & d_3 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix} \det \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix} \cdots \end{pmatrix}$$

$$= \begin{pmatrix} \det \mathcal{M} & 0 & 0 & \cdots \\ 0 & \det \mathcal{M} & 0 & \cdots \\ 0 & 0 & \det \mathcal{M} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}. \tag{B.8c}$$

In the last line we use the fact that any determinant with two identical rows must vanish. Now the general formula for the inverse $\mathcal{M}^{-1}$ follows if we

can divide Eq. (B.8) by det $\mathcal{M}$:

$$\mathcal{M}^{-1} = \mathcal{M}^{\text{adj}}/(\det \mathcal{M}). \tag{B.9}$$

This is the case when det $\mathcal{M} \neq 0$. Then from Eq. (B.3) we may write the general solution to the linear equation $\mathcal{M}x = y$, which is called KRAMER'S RULE:

$$x = \mathcal{M}^{-1}y = \frac{\mathcal{M}^{\text{adj}}}{\det \mathcal{M}}y = \begin{pmatrix} \det \begin{vmatrix} y_1 & a_2 & a_3 & a_4 & \cdots \\ y_2 & b_2 & b_3 & b_4 & \cdots \\ y_3 & c_2 & c_3 & c_4 & \cdots \\ y_4 & d_2 & d_3 & d_4 & \cdots \end{vmatrix} \\ \det \begin{vmatrix} a_1 & y_1 & a_3 & b_4 & \cdots \\ b_1 & y_2 & b_3 & b_4 & \cdots \\ c_1 & y_3 & c_3 & c_4 & \cdots \\ d_1 & y_4 & d_3 & d_4 & \cdots \end{vmatrix} \\ \det \begin{vmatrix} a_1 & a_2 & y_1 & a_4 & \cdots \\ b_1 & b_2 & y_2 & b_4 & \cdots \\ c_1 & c_2 & y_3 & c_4 & \cdots \\ d_1 & d_2 & y_4 & d_4 & \cdots \end{vmatrix} \end{pmatrix} \cdot (\det \mathcal{M})^{-1}. \tag{B.10}$$

A linear equation with $y = 0$ is called a HOMOGENEOUS equation. Equation (B.11) is an example:

$$\mathcal{M} \cdot x = 0 \tag{B.11}$$

Now if det $\mathcal{M} \neq 0$ then according to Kramer's rule the only solutions are zero vectors $x = 0$. However, if det $\mathcal{M} = 0$ there will exist nonzero solutions as shown in Chapter 1, Section 1.2.B. A matrix $\mathcal{M}$ is said to be SINGULAR if det $\mathcal{M}$ is zero and NONSINGULAR if det $\mathcal{M}$ is nonzero.

In Appendix A, a set of vectors $|a\rangle, |b\rangle, |c\rangle, \ldots$ were defined to be linearly dependent if and only if a relation of the form of Eq. (B.12) could exist for values of the coefficients $\alpha, \beta, \gamma, \ldots$ not all zero:

$$\alpha|a\rangle + \beta|b\rangle + \gamma|c\rangle + \cdots = 0. \tag{B.12}$$

By taking the scalar product of this relation with each vector in turn, one

derives a simple matrix equation which can be used to test for linear dependence.

$$\begin{aligned}
\alpha\langle a|a\rangle + \beta\langle a|b\rangle + \gamma\langle a|c\rangle + \cdots &= 0,\\
\alpha\langle b|a\rangle + \beta\langle b|b\rangle + \gamma\langle b|c\rangle + \cdots &= 0,\\
\alpha\langle c|a\rangle + \beta\langle c|b\rangle + \gamma\langle c|c\rangle + \cdots &= 0,\\
\begin{pmatrix} \langle a|a\rangle & \langle a|b\rangle & \langle a|c\rangle & \cdots \\ \langle b|a\rangle & \langle b|b\rangle & \langle b|c\rangle & \cdots \\ \langle c|a\rangle & \langle c|b\rangle & \langle c|c\rangle & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}
\begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \vdots \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}.
\end{aligned}$$

Clearly nonzero $\alpha, \beta, \gamma, \ldots$ exist only if the determinant of the matrix $\langle i|j\rangle$ is zero.

APPENDIX C. PROOF THAT HERMITIAN AND UNITARY MATRICES ARE DIAGONALIZABLE

If a matrix $\mathscr{M}$ satisfied a minimal equation (MEq) with no repeated roots r_j, then we showed in Section 1.2.B(a) how a set of relatively prime polynomials $\mathscr{P}_{r_j}$ of this matrix could be made to obey the completeness and orthonormality relations of Eqs. (1.2.17) and (1.2.18). This amounted to giving a complete set of eigenvectors for the matrix, and guaranteed that $\mathscr{M}$ was DIAGONALIZABLE, or spectrally decomposable.

Now suppose instead that a repeated root r_j did appear in the MEq, as shown in Eq. (C.1):

$$(\mathscr{M} - r_1 1)\cdots(\mathscr{M} - r_j 1)(\mathscr{M} - r_j 1)\cdots(\mathscr{M} - r_n 1) = 0. \qquad \text{(C.1)}$$

This implies that matrix $\mathscr{N}$ in Eq. (C.2), made from the minimal polynomial by lifting just one redundant factor $(\mathscr{M} - r_j 1)$, must be a nonzero matrix whose square is zero (such a matrix is called a NILPOTENT):

$$\mathscr{N} = (\mathscr{M} - r_1 1)\cdots(1)(\mathscr{M} - r_j 1)\cdots(\mathscr{M} - r_n 1) \neq 0, \qquad \text{(C.2a)}$$

$$\mathscr{N}\mathscr{N} = 0. \qquad \text{(C.2b)}$$

Now, this is impossible if $\mathscr{M} \equiv \mathscr{H} \equiv \mathscr{H}^\dagger$ is Hermitian, since then (we use here the fact that all roots of $\mathscr{H} = \mathscr{H}^\dagger$ must be real... see Problem 1.2.7) $\mathscr{N}$ must also be Hermitian; Eq. (C.2b) contradicts Eq. (C.2a) if $\mathscr{N}_{ij} = \mathscr{N}_{ji}^*$. As seen in Eq. (C.3), we must have

$$(\mathscr{N}\mathscr{N})_{ii} = 0 = \sum_j \mathscr{N}_{ij}\mathscr{N}_{ji} = \sum_j \mathscr{N}_{ij}\mathscr{N}_{ij}^* = \sum_j |\mathscr{N}_{ij}|^2 \qquad \text{(C.3)}$$

Hence a Hermitian matrix must be diagonalizable, i.e., decomposable, as in Eq. (C.4a) to a sum of idempotents $\mathscr{P}_{r_i}$ which satisfy completeness relations [Eq. (C.4b)]. By construction [recall Eq. (1.2.12)] the $\mathscr{P}_{r_i}$ are Hermitian [Eq. (C.4c)]:

$$\mathscr{H} = \sum_{r_i} r_i \mathscr{P}_{r_i} \qquad (r_i^* = r_i), \tag{C.4a}$$

$$1 = \sum_{r_i} \mathscr{P}_{r_i}, \tag{C.4b}$$

$$\mathscr{P}_{r_i}^\dagger = \mathscr{P}_{r_i} = \prod_{r_l \neq r_i} (\mathscr{H} - r_l 1) \Big/ \prod_{r_l \neq r_i} (r_i - r_l). \tag{C.4c}$$

The same applies to any ANTI-HERMITIAN matrix $\mathscr{M} = \mathscr{A} = -\mathscr{A}^\dagger$, since $i\mathscr{A}$ is Hermitian, as given in Eqs. (C.5). Note that $\mathscr{A}$ eigenvalues are purely imaginary:

$$\mathscr{A} = \sum_{a_j} a_j \mathscr{P}_{a_j} \qquad (a_j^* = -a_j), \tag{C.5a}$$

$$1 = \sum_{a_j} \mathscr{P}_{a_j}, \tag{C.5b}$$

$$\mathscr{P}_{a_j}^\dagger = \mathscr{P}_{a_j}. \tag{C.5c}$$

Finally, we can see that the same can be done for a unitary matrix $\mathscr{U}^\dagger \mathscr{U} = 1 = \mathscr{U}\mathscr{U}^\dagger$. First observe that the Hermitian and anti-Hermitian matrices $\mathscr{H}$ and $\mathscr{A}$, made from $\mathscr{U}$ in Eq. (C.6), will commute ($\mathscr{H}\mathscr{A} = \mathscr{A}\mathscr{H}$):

$$\mathscr{H} = \mathscr{U} + \mathscr{U}^\dagger, \qquad \mathscr{A} = \mathscr{U} - \mathscr{U}^\dagger. \tag{C.6}$$

Therefore, the product of the completeness relations of Eqs. (C.5b) and (C.4b) must yield a new set of idempotents $\mathscr{P}_{r_i, a_j} = \mathscr{P}_{r_i}\mathscr{P}_{a_j}$, which can simultaneously decompose $\mathscr{H}$, $\mathscr{A}$, and any combinations thereof, including $\mathscr{U} = (\mathscr{H} + \mathscr{A})/2$.

$$1 \cdot 1 = \left(\sum_{r_i} \mathscr{P}_{r_i}\right)\left(\sum_{a_j} \mathscr{P}_{a_j}\right) = \sum_{r_i a_j} \mathscr{P}_{(r_i a_j)}, \tag{C.7a}$$

$$\mathscr{U} = \sum_{r_i a_j} u_{ij} \mathscr{P}_{(r_i a_j)}, \qquad u_{ij} = (r_i + a_j)/2 = \frac{1}{u_{ij}^*}, \tag{C.7b}$$

$$\mathscr{P}_{(r_i a_j)}^\dagger = \mathscr{P}_{(r_i a_j)}. \tag{C.7c}$$

(Many of the products $\mathscr{P}_{r_i}\mathscr{P}_{a_j}$ may be zero, but whatever is left must be complete.) The eigenvalues of $\mathscr{U}$ are of the form $u = e^{i\phi}$ if $\mathscr{U}$ is unitary.

ADDITIONAL READING

One of the best introductions to the fundamental quantum theory is found in the third and final volume of the *Feynman Lectures*. One should read all three volumes, but the final volume is quite self-contained and may be worth the price of the whole set.

R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1967), Volumes I–III).

Feynman's unusual approach to quantum theory introduces the Dirac bra-ket transamplitudes $\langle b|a\rangle$ and their matrix mechanics in a physical way. Two- or three-state systems are used as examples in the primary development. Then infinite-dimensional amplitudes or wave functions $\langle x|\Psi\rangle = \Psi(x)$ and differential operators are treated subsequently.

Another text which begins by discussing amplitude mechanics is by Gordan Baym.

G. Baym, *Lectures on Quantum Mechanics* (Benjamin Advanced Book Program, Reading, MA, 1973).

Interpretation of quantum states is debated in *Physics Today* (April 1993, page 13).

A discussion of spectral decomposition of matrices and the algebra of projection operators should be found in any good book on mathematical methods for physics or chemistry. However, most books written for physicists do not cover this subject very well, if at all. One notable exception which has recently been published is a text by Hasani.

Sadri Hasani, *Foundations of Mathematical Physics* (Allyn and Bacon, Boston, 1991).

This text also describes the spectral decomposition and resolvants of differential operators.

It is easier to find treatments of matrix spectral decomposition in the mathematical literature. An excellent readable treatment is found on p. 155 of the following older but well known text. It features the spectral theorem as the main result of the book.

Paul R. Halmos, *Finite Dimensional Vector Spaces* (Van Nostrand, Princeton, 1958).

PROBLEMS

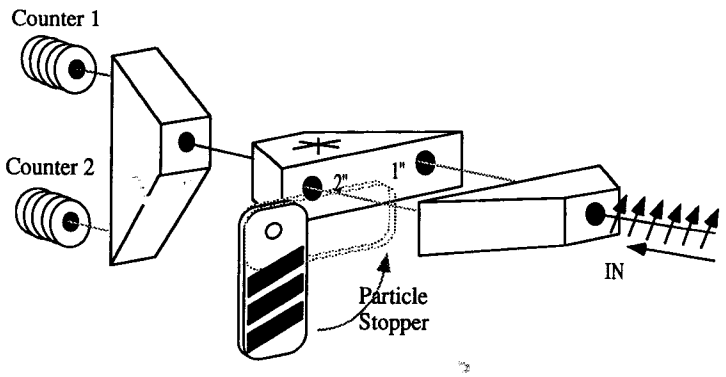
Section 1.1

1.1.1 (Beam stoppers) Consider an incoming beam of spin-$\frac{1}{2}$ particles approaching the apparatus shown in the following diagram. A "stopper" can block the intermediate (2)" beam if inserted as shown in the diagram.

Use the following notation for the states having different analyzer angle θ.

$$|1\rangle = |\text{up}\rangle \quad |1'\rangle = |\text{N}\rangle \quad |1''\rangle = |\text{right}\rangle$$
$$|2\rangle = |\text{dn}\rangle \quad |2'\rangle = |\text{S}\rangle \quad |2''\rangle = |\text{left}\rangle$$
$$\theta = 0, \quad 0 < \theta < \pi/2, \quad \theta = \pi/2.$$

Compute the percentage of particles that get to counter (1)...	without stopper	with stopper
for an incoming beam polarized in state: $\|1\rangle$	_____	_____
... in state: $\|2\rangle$	_____	_____
... in state: $\|1'\rangle$	_____	_____

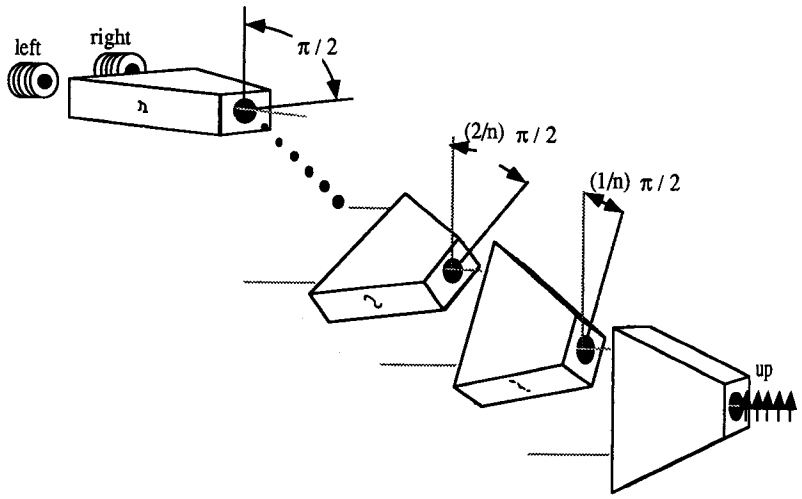


1.1.2 (Pointing spins the "right" way)

(a) How many of the spins that were originally up ($|1\rangle$) going into this apparatus* will end up "right"? Give answer as a function of n and see how it behaves for $n = 1, 2, 3, \ldots, \infty$.

(b) Can you design an apparatus* that duplicates the effect of the elementary operator $\mathbf{e}_{22}$? How about $\mathbf{e}_{12}$?

*"Apparatus" means some arrangement of Stern-Gerlach analyzers and stoppers.



1.1.3 **(a)** Prove that $(AB)^\dagger = B^\dagger A^\dagger$ using components A_{ij} and B_{ij} of $n \times n$ matrices A and B.

(b) Prove that $(A|x\rangle)^\dagger = \langle x|A^\dagger$.

(c) Expand to find $(ABCD)^\dagger = ?, (\langle x|A|y\rangle)^\dagger = ?, (|x\rangle\langle y|)^\dagger ?, (\mathbf{e}_{12})^\dagger ?$

Section 1.2

1.2.1 (Back to your roots)

For the matrix $M = \begin{pmatrix} 2 & -1 \\ -6 & 1 \end{pmatrix}$ find

(a) Eigenvalues of M.

(b) Spectral decomposition of M.

(c) The bra and ket eigenvectors of M.

(d) All the square roots of M: $\sqrt{M} = \sqrt{\begin{pmatrix} 2 & -1 \\ -6 & 1 \end{pmatrix}}$

1.2.2 Do the same as in Problem 1.2.1 for the matrix $M = \begin{pmatrix} 11 & 7 & -9 \\ 7 & 11 & -9 \\ -9 & -9 & 27 \end{pmatrix}$.

1.2.3 (Secular behavior)

The polynomial form for the secular equation of a general $n \times n$ matrix M is

$$\begin{aligned} S(\varepsilon) &= \det|M - \varepsilon 1| \\ &= \varepsilon^n + a_1\varepsilon^{n-1} + a_2\varepsilon^{n-2} + \cdots + a_{n-1}\varepsilon + a_0 = 0. \end{aligned}$$

(a) Derive general formulas for the coefficients $a_1, a_2, a_3, \ldots, a_{n-1}, a_n$ in terms of the eigenvalues $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n\}$ of matrix M.

(b) Derive general formulas for the coefficients $a_1, a_2, a_3, \ldots, a_{n-1}, a_n$ in terms of the components $\{M_{ij}\}$ of matrix M. (Hint: Use the ε-tensor definition

$$\det|M| = \sum_{i_1 i_2 \cdots i_m} \varepsilon_{i_1 i_2 \cdots i_m} M_{1,i_1} M_{2,i_2} \cdots M_{n,i_n}$$

for determinants and expand the determinantal secular expression.)

(c) Do the coefficients a_j change if there is a change of basis?

[*Note: The results of (a) and (b) are particularly important and will be helpful for later theory and problems throughout the book.*]

1.2.4 (Multiple degeneracy) The matrix $\rho = \rho \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ represents a simple example of what is called a *pairing* operator in nuclear and superconductivity theory. It has number of repeated or degenerate eigenvalues.

(a) Find the secular equation, the Hamiltonian Cayley equation, and the minimal equation of ρ. (*Hint: Problem 1.2.3 is useful.*) Find the eigenvalues.

(b) Compute the projection operators for each distinct eigenvalue and write the spectral decomposition for this matrix.

(c) Find all the square roots $\sqrt{\rho}$ of matrix ρ.

1.2.5 (Knowing spectral decomposition backwards and forwards)

(a) Use what we have discussed about spectral decomposition and $\otimes$ products to do an eigenvalue problem backwards. Find the matrix $M = M^\dagger$ which has the following eigensolutions:

$$M\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad M\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}, \quad M\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

(b) Find M^{100}.

1.2.6 (Commuting observables)

(a) Use the techniques in Chapter 1 (Section 1.2Bd) to find a single set of projection operators which spectrally decompose both matrices:

$$M = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}.$$

(b) Use the result of (a) to find a single transformation T which diagonalizes M and N.

1.2.7 What are the conjugation relations (if any) for eigenvalues (E_j *and* E_j^*), projection operators ($\mathbf{P}_{Ej}$ and $\mathbf{P}_{Ej}^\dagger$), and eigenvectors ($|E_j\rangle$ and $\langle E_j|$) in the cases that operators are (a) Hermitian: $\mathbf{H} = \mathbf{H}^\dagger$, (b) anti-Hermitian: $\mathbf{A}^\dagger = -\mathbf{A}$, or (c) unitary: $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. Check your conclusions by spectrally decomposing and diagonalizing the following examples:

(a) $\langle H \rangle = \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ i/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$. **(b)** $\langle A \rangle = \begin{pmatrix} 0 & i/\sqrt{2} \\ i/\sqrt{2} & 0 \end{pmatrix}$.

(c) $\langle U \rangle = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{pmatrix}$.

1.2.8 Does $M^\dagger M = 1$ imply that $MM^\dagger = 1$, as well?

(a) Prove or disprove in the case that M is a finite $n \times n$ matrix.

(b) What if M is an *infinite* matrix? (Examples of infinite matrices are the representations of creation and destruction operators $\mathbf{a}$ and $\mathbf{a}^\dagger$ in the quantum harmonic oscillator eigenstates $\{|0\rangle, |1\rangle = \mathbf{a}^\dagger|0\rangle, |2\rangle = \mathbf{a}^\dagger|1\rangle, \ldots\}$.)