Chapter 8

Developments

We have now constructed the basic framework of quantum mechanics. In this chapter, we will introduce some secondary concepts that are important in practice.

8.1 Matrix elements

A general algebraic element, say one representing a physical property, can be written as a linear combination of the n^2 measurement symbols,

$$|a'a''| = |a'\rangle\langle a''|. \tag{8.1}$$

We see this as follows:

$$\begin{aligned} X &= 1X1 = \left(\sum_{a'} |a'\rangle\langle a'|\right) X \left(\sum_{a''} |a''\rangle\langle a''|\right) \\ &= \sum_{a'a''} |a'\rangle\langle a'|X|a''\rangle\langle a''| = \sum_{a'a''} \langle a'|X|a''\rangle |a'a''|, \end{aligned} \tag{8.2}$$

where we see the coefficients of expansion, $\langle a'|X|a''\rangle$, of X in terms of the measurment symbols $|a'a''|$. We call $\langle a'|X|a''\rangle$ the a'-a'' matrix element of X. This generalizes what we had in Eqs. (7.37a) and (7.37b) for the σ's in terms of a systematic notation. We write these $n^2 = n \times n$ elements in an array: $\langle a'|X|a''\rangle$ is the element in the a'th row, a''th column:

$$X = \begin{pmatrix} \langle a_1|X|a_1\rangle & \langle a_1|X|a_2\rangle & \dots & \langle a_1|X|a_n\rangle \\ \langle a_2|X|a_1\rangle & \langle a_2|X|a_2\rangle & \dots & \langle a_2|X|a_n\rangle \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \langle a_n|X|a_1\rangle & \langle a_n|X|a_2\rangle & \dots & \langle a_n|X|a_n\rangle \end{pmatrix}. \tag{8.3}$$

Instead of working with the measurement symbol representation of X, we can work directly with this numerical coefficient array, the matrix representing X.

What are the algebraic properties of these arrays? They correspond to the algebraic properties of the elements of our algebra:

$$\text{addition:} \quad X + Y, \tag{8.4a}$$
$$\text{multiplication by numbers:} \quad \lambda X, \tag{8.4b}$$
$$\text{multiplication:} \quad XY. \tag{8.4c}$$

For addition, since

$$X = \sum_{a'a''} \langle a'|X|a''\rangle |a'a''|, \tag{8.5a}$$
$$Y = \sum_{a'a''} \langle a'|Y|a''\rangle |a'a''|, \tag{8.5b}$$

we have

$$\begin{aligned} X + Y &= \sum_{a'a''} \left(\langle a'|X|a''\rangle + \langle a'|Y|a''\rangle\right) |a'a''| \\ &= \sum_{a'a''} \langle a'|X + Y|a''\rangle |a'a''|, \end{aligned} \tag{8.6}$$

where the last line follows from the definition of the matrix element of $X + Y$. An array for $X + Y$ is obtained by adding corresponding coefficients in arrays for X and Y (same row, same column). For example,

$$\sigma_x + \sigma_y = \begin{pmatrix} 0 & 1-i \\ 1+i & 0 \end{pmatrix}. \tag{8.7}$$

This result may also be seen from the distributive property of multiplication:

$$\langle a'|X + Y|a''\rangle = \langle a'| \left(X|a''\rangle + Y|a''\rangle\right) = \langle a'|X|a''\rangle + \langle a'|Y|a''\rangle. \tag{8.8}$$

For multiplication by a number,

$$\lambda X = \sum_{a'a''} \lambda \langle a'|X|a''\rangle |a'a''| = \sum_{a'a''} \langle a'|\lambda X|a''\rangle |a'a''|, \tag{8.9}$$

again from the definition. So, for example,

$$2\sigma_x = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}. \tag{8.10}$$

Not so trivial is multiplication:

$$\begin{aligned} XY &= \sum_{a'a''} \langle a'|X|a''\rangle |a'a''| \sum_{a'''a^{iv}} \langle a'''|Y|a^{iv}\rangle |a'''a^{iv}| \\ &= \sum_{a'a''a'''a^{iv}} \langle a'|X|a''\rangle \langle a'''|Y|a^{iv}\rangle |a'a''||a'''a^{iv}|, \end{aligned} \tag{8.11}$$

so since $|a'a''||a'''a^{iv}| = \delta(a'', a''')|a'a^{iv}|$,

$$XY = \sum_{a'a'''} \langle a'|XY|a'''\rangle |a'a'''|, \tag{8.12}$$

where

$$\langle a'|XY|a'''\rangle = \sum_{a''} \langle a'|X|a''\rangle\langle a''|Y|a'''\rangle. \tag{8.13}$$

A second derivation uses the expression of a completely nonselective measurement:

$$1 = \sum_{a''} |a''\rangle\langle a''|, \tag{8.14}$$

so

$$\langle a'|XY|a'''\rangle = \langle a'|X1Y|a'''\rangle = \sum_{a''} \langle a'|X|a''\rangle\langle a''|Y|a'''\rangle. \tag{8.15}$$

In words, the rule is: To find the entry of XY is the a'th row, a'''th column, take in entries in the a'th row of X and in the a'''th column of Y, multiply them together term by term, and add the results together. This row on column multiplication rule is indeed that of *matrices*.

Here is an example:

$$\sigma_x \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\sigma_z. \tag{8.16}$$

To reiterate, we call $\langle a'|X|a''\rangle$ an element of the matrix representing X, or in short, the *matrix element* of X (between the a' a'' states)

How do we represent the multiplication of algebraic elements and vectors, such as

$$\sigma_x|+\rangle = |-\rangle? \tag{8.17}$$

In the a "coordinate system" we ask, what are the components of the vector $X|\,\rangle$, in terms of the components of $|\,\rangle$?

$$\langle a'|X|\,\rangle = \sum_{a''} \langle a'|X|a''\rangle\langle a''|\,\rangle, \tag{8.18}$$

which represents row on column multiplication of the matrix for X on the column vector, the wavefunction, representing $|\,\rangle$, $\psi(a'') = \langle a''|\,\rangle$.

Here's again the spin-1/2 example (8.17), which is represented by

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{8.19}$$

because $\langle +|+\rangle = 1$, $\langle -|+\rangle = 0$. We can also multiply to the left:

$$\langle -|\sigma_x = \langle +|, \tag{8.20}$$

is represented by

$$(0, 1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (1, 0). \tag{8.21}$$

The general rule in the latter case is

$$\langle\,|X|a''\rangle = \sum_{a'} \langle\,|a'\rangle\langle a'|X|a''\rangle = \sum_{a'} \psi(a')^* \langle a'|X|a''\rangle, \tag{8.22}$$

again, row on column multiplication.

8.2 Eigenvectors and eigenvalues

Consider the physical quantity A itself, which has possible values a':

$$A = \sum_{a'} a'|a'| = \sum_{a'} a'|a'a'|. \tag{8.23}$$

The matrix of A is thus very special,

$$\langle a'|A|a''\rangle = a'\delta(a', a''), \tag{8.24}$$

or

$$A = \begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & a_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_n \end{pmatrix}, \tag{8.25}$$

which is a diagonal matrix. For example, σ_z is diagonal, for it has a definite value in the states $|+\rangle$ and $|-\rangle$. Here

$$A|a'\rangle = a'|a'\rangle, \quad \langle a'|A = \langle a'|a', \tag{8.26}$$

that is $|a'\rangle$, $\langle a'|$ represent a state in which A has the value a'.

Before proceeding, let's introduce a bit more language. Since an element of the measurment algebra X acts on a vector to produce another vector,

$$X|a'\rangle, \tag{8.27}$$

we call X an *operator*, since it operates on vectors to produce vectors.

Now when an operator acts on a vector to give back the *same* vector, apart from a numerical multiple,

$$A|a'\rangle = a'|a'\rangle, \tag{8.28}$$

we say that $|a'\rangle$ is an *eigenvector* ("characteristic" or "proper" vector) of A, and the number a' is an *eigenvalue* (or "characteristic" value) of A. An eigenvector represents a state in which the operator has a definite value, the eigenvalue. The construction

$$A = \sum_{a'} a'|a'\rangle\langle a'| \tag{8.29}$$

expands A in terms of its eigenvectors and eigenvalues.

We can now ask a mathematical question. Given a certain operator, what are its eigenvectors and eigenvalues? The corresponding physical question is: What are the states in which the physical property has a definite value, and what are those values?

8.2.1 Eigenvectors for spin-1/2

Consider

$$\sigma_{z'} = \sigma_x \sin\theta\cos\phi + \sigma_y \sin\theta\sin\phi + \sigma_z\cos\theta. \tag{8.30}$$

We ask, what are the possible value of $\sigma_{z'}$, and what are the states in which $\sigma_{z'}$ assumes those definite values? To find the eigenvalues and eigenvectors of $\sigma_{z'}$ it is convenient to work in the matrix representation, where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.31}$$

Thus

$$\sigma_{z'} = \begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix}. \tag{8.32}$$

The eigenvalue equation

$$\sigma_{z'}|\sigma'_{z'}\rangle = \sigma'_{z'}|\sigma'_{z'}\rangle \tag{8.33}$$

corresponds to the matrix equation

$$\begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix} \begin{pmatrix} \psi(+) \\ \psi(-) \end{pmatrix} = \sigma' \begin{pmatrix} \psi(+) \\ \psi(-) \end{pmatrix}, \tag{8.34}$$

where we have simplified the notation by writing $\sigma'_{z'} = \sigma'$. Reference is here being implicitly made to a standard set of vectors, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, corresponding to definite values of σ_z, ± 1, respectively. Rewrite this equation as

$$\begin{pmatrix} \cos\theta - \sigma' & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta - \sigma' \end{pmatrix} \begin{pmatrix} \psi(+) \\ \psi(-) \end{pmatrix} = 0. \tag{8.35}$$

This system of two simultaneous homogeneous equations has a nonzero solution for ψ only if the coefficient matrix has a vanishing determinant:

$$\begin{aligned} 0 &= \det \begin{pmatrix} \cos\theta - \sigma' & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta - \sigma' \end{pmatrix} \\ &= \sigma'^2 - \cos^2\theta - \sin^2\theta = \sigma'^2 - 1, \end{aligned} \tag{8.36}$$

so as we well know,

$$\sigma'^2 = 1, \quad \sigma' = \pm 1. \tag{8.37}$$

What are the corresponding eigenvectors (wavefunctions)? Write out the equations explicitly for $\sigma' = +1$:

$$(\cos\theta - 1)\psi(+) + \sin\theta\, e^{-i\phi}\psi(-) = 0, \tag{8.38a}$$
$$\sin\theta\, e^{i\phi}\psi(+) - (\cos\theta + 1)\psi(-) = 0, \tag{8.38b}$$

or introducing half-angles,

$$\sin^2\frac{\theta}{2}\,\psi(+) = \sin\frac{\theta}{2}\cos\frac{\theta}{2}\, e^{-i\phi}\psi(-), \tag{8.39a}$$
$$\sin\frac{\theta}{2}\cos\frac{\theta}{2}\, e^{i\phi}\psi(+) = \cos^2\frac{\theta}{2}\,\psi(-). \tag{8.39b}$$

Both these equations are equivalent to

$$\sin\frac{\theta}{2}\,\psi(+) = \cos\frac{\theta}{2}\,e^{-i\phi}\psi(-). \tag{8.40}$$

The two equations give the same information because of the determinant condition—we can only extract two pieces of information, here σ' and $\psi(+)/\psi(-)$, from two equations. The above equation will be satisfied if

$$\psi(+) = A\cos\frac{\theta}{2}e^{-i\phi/2}, \quad \psi(-) = A\sin\frac{\theta}{2}e^{i\phi/2}. \tag{8.41}$$

Evidently, the eigenvector equation does not determine an overall factor A multiplying both components of a wavefunction. A is nearly determined by the requirement that a physical state be described by a unit vector. In terms of the wavefunction this means

$$|\psi(+)|^2 + |\psi(-)|^2 = 1, \tag{8.42}$$

or

$$|A|^2\left(\cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}\right) = 1, \tag{8.43}$$

or

$$|A| = 1, \quad A = e^{i\alpha}, \tag{8.44}$$

where α is an arbitrary real number. Thus

$$\psi_{+z'} = e^{i\alpha}\begin{pmatrix}\cos\frac{\theta}{2}e^{-i\phi/2}\\ \sin\frac{\theta}{2}e^{i\phi/2}\end{pmatrix}; \tag{8.45}$$

when $\alpha = 0$, this agrees with the wavefunction found in Eq. (7.36); when $\alpha = \phi/2$, this agrees with the wavefunction found in Eq. (4.35).

For $\sigma' = -1$, we'll look at just one of the two equivalent equations:

$$(\cos\theta + 1)\psi(+) + \sin\theta\, e^{-i\phi}\psi(-) = 0, \tag{8.46}$$

or

$$\cos^2\frac{\theta}{2}\,\psi(+) + \cos\frac{\theta}{2}\sin\frac{\theta}{2}\,e^{-i\phi}\psi(-) = 0, \tag{8.47}$$

or equivalently

$$\cos\frac{\theta}{2}\psi(+) = -\sin\frac{\theta}{2}e^{-i\phi}\psi(-), \tag{8.48}$$

which determines the wavefunction up to a multiplicative factor,

$$\psi(+) = -B\sin\frac{\theta}{2}e^{-i\phi/2}, \quad \psi(-) = B\cos\frac{\theta}{2}e^{i\phi/2}, \tag{8.49}$$

and again the normalization

$$|\psi(+)|^2 + |\psi(-)|^2 = 1 \tag{8.50}$$

implies, for β real,

$$|B| = 1, \quad \text{or} \quad B = e^{i\beta}. \tag{8.51}$$

Thus the second eigenvector, belonging to $\sigma'_{z'} = -1$ is

$$\psi_{-z'} = e^{i\beta} \begin{pmatrix} -\sin\frac{\theta}{2} e^{-i\phi/2} \\ \cos\frac{\theta}{2} e^{i\phi/2} \end{pmatrix}, \tag{8.52}$$

which again reproduces previous results.

The orthogonality of the two different wavefunctions, $\psi_{+z'}$, $\psi_{-z'}$, follows automatically from the eigenvalue problem: Multiply the latter, in general,

$$A|a''\rangle = a''|a''\rangle, \tag{8.53}$$

on the left by $\langle a'|$ to get

$$\langle a'|A|a''\rangle = \langle a'|a''|a''\rangle = a''\langle a'|a''\rangle, \tag{8.54}$$

while the corresponding eigenvalue problem for left vectors,

$$\langle a'|A = a'\langle a'|, \tag{8.55}$$

may be multiplied on the right by $|a''\rangle$:

$$\langle a'|A|a''\rangle = a'\langle a'|a''\rangle. \tag{8.56}$$

Eqs. (8.54) and (8.56) are the same: in the first, we consider A as acting to the right, in the second to the left. In either case, we have the identical matrix element of A. Thus we conclude

$$(a' - a'')\langle a'|a''\rangle = 0. \tag{8.57}$$

So if $a' \neq a''$, $\langle a'|a''\rangle = 0$. Eigenvectors belonging to different eigenvalues are orthogonal. In other words, vectors belonging to physically different states are orthogonal.

The following sketches a general procedure. Suppose we solve the eigenvalue problem

$$A|a'\rangle = a'|a'\rangle, \tag{8.58}$$

and find all n eigenvalues and eigenvectors. Suppose that all the a''s are different. Then the n different eigenvectors $|a'\rangle$ are orthogonal to each other. These vectors can be normalized so we obtain a set of n orthonormal vectors—these vectors form a coordinate system since there are only n indepdendent vectors in the n-dimensional space of states.

$$\langle a'|a''\rangle = \delta(a', a''). \tag{8.59}$$

Any vector in the space can be expressed as a linear combination of these vectors, which means

$$1 = \sum_{a'} |a'\rangle\langle a'|. \tag{8.60}$$

We can then construct A in terms of its eigenvectors and eigenvalues,

$$A = A1 = A\sum_{a'} |a'\rangle\langle a'| = \sum_{a'} a'|a'\rangle\langle a'|. \tag{8.61}$$

Let's now eliminate a stage of this process and get the normalization automatically. (We had done this before, even before we constructed the machinery of quantum mechanics.) We do this by considering the measurement symbol, in the spin-1/2 example:

$$|\sigma'_{z'} = +1| = \frac{1+\sigma_{z'}}{2} = |+z'\rangle\langle +z'|, \tag{8.62}$$

or, in terms of wavefunctions and matrices,

$$\langle\sigma'_z| + z'\rangle\langle +z'|\sigma''_z\rangle = \psi_{+z'}(\sigma'_z)\psi_{+z'}(\sigma''_z)^* = \langle\sigma'_z|\frac{1+\sigma_{z'}}{2}|\sigma''_z\rangle. \tag{8.63}$$

Explicitly, since as a matrix

$$\sigma_{z'} = \langle\sigma'_z|\sigma_{z'}|\sigma''_z\rangle = \begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix}, \tag{8.64}$$

we have as a matrix

$$\begin{aligned} \frac{1+\sigma_{z'}}{2} &= \langle\sigma'_z|\frac{1+\sigma_{z'}}{2}|\sigma''_z\rangle = \begin{pmatrix} \cos^2\frac{\theta}{2} & \cos\frac{\theta}{2}\sin\frac{\theta}{2}\, e^{-i\phi} \\ \cos\frac{\theta}{2}\sin\frac{\theta}{2}\, e^{i\phi} & \sin^2\frac{\theta}{2} \end{pmatrix} \\ &= \begin{pmatrix} \cos\frac{\theta}{2}e^{-i\phi/2} \\ \sin\frac{\theta}{2}e^{i\phi/2} \end{pmatrix} \left(\cos\frac{\theta}{2}e^{i\phi/2}, \sin\frac{\theta}{2}e^{-i\phi/2}\right), \end{aligned} \tag{8.65}$$

where we have noted that row on column multiplication of a column vector on a row vector gives a (2×2) matrix. The two factors are $\psi_{+z'}$ and $\psi^*_{+z'}$. Note that three of the matrix elements determine the coefficients in the wavefunctions; the fourth provides a test of consistency. A second test is passed when we see that the row and column vectors here are complex conjugates of each other. There is a phase ambiguity in this separation—we can multiply the column vector by $e^{i\alpha}$, and the row vector by $e^{-i\alpha}$. Thus the wavefunction is

$$\psi_{z'} = e^{i\alpha}\begin{pmatrix} \cos\frac{\theta}{2}e^{-i\phi/2} \\ \sin\frac{\theta}{2}e^{i\phi/2} \end{pmatrix}, \tag{8.66}$$

which is the same as Eq. (8.45).

Similarly,

$$\begin{aligned} \frac{1-\sigma_{z'}}{2} &= \begin{pmatrix} \sin^2\frac{\theta}{2} & -\sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{-i\phi} \\ -\sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{i\phi} & \cos^2\frac{\theta}{2} \end{pmatrix} \\ &= \begin{pmatrix} -\sin\frac{\theta}{2}e^{-i\phi/2} \\ \cos\frac{\theta}{2}e^{i\phi/2} \end{pmatrix} \left(-\sin\frac{\theta}{2}e^{i\phi/2}, \cos\frac{\theta}{2}e^{-i\phi/2}\right) \\ &= \psi_{-z'}\psi^*_{-z'}, \end{aligned} \tag{8.67}$$

where

$$\psi_{-z'} = e^{i\beta}\begin{pmatrix} -\sin\frac{\theta}{2}e^{-i\phi/2} \\ \cos\frac{\theta}{2}e^{i\phi/2} \end{pmatrix}, \tag{8.68}$$

which is the same as Eq. (8.52).

8.3 The adjoint

From time to time we have mentioned our freedom to read equations from left to right or from right to left. Suppose we have

$$|a'a''|. \tag{8.69}$$

This means:

- L→R: select a', put into a''.
- R→L: select a'', put into a'.

We now invent an operation to express this change in the convention in terms of which we read the symbols. We'll call this operation the *adjoint*, denoted by † (dagger), defined by

$$|a'a''|^\dagger = |a''a'|; \tag{8.70}$$

† means reverse the convention of reading symbols, but write the result in the original convention. Thus, in the L→ R convention,

- $|a'a''|$ means: select only a', put into a'',
- $|a'a''|^\dagger = |a''a'|$ means: select only a'', put into a'.

Suppose we had a sequence of measurements:

$$\left(|a'a''||a'''a^{iv}|\right)^\dagger = |a^{iv}a'''||a''a'|, \tag{8.71}$$

since † means read everything the other way. But the last expression is the same as

$$|a'''a^{iv}|^\dagger|a'a''|^\dagger. \tag{8.72}$$

So, the adjoint of a product is the product of the adjoints *in the opposite order*. Since any operator is a linear combination of measurement symbols of the above type, and the adjoint operation has no effect on addition (the order of addition is irrelevant), we have the general algebraic statement

$$(XY)^\dagger = Y^\dagger X^\dagger. \tag{8.73}$$

Suppose we have

$$|a'b'||c'd'|, \tag{8.74}$$

where a', b', c', d' refer to different physical properties (e.g., σ_x, σ_y, ...). Then the adjoint of this is

$$(|a'b'||c'd'|)^\dagger = |d'c'||b'a'| = |c'd'|^\dagger|a'b'|^\dagger. \tag{8.75}$$

But

$$|a'b'||c'd'| = |a'\rangle\langle b'|c'\rangle\langle d'|, \tag{8.76}$$

while the adjoint is

$$(|a'b'||c'd'|)^\dagger = |d'\rangle\langle c'|b'\rangle\langle a'| = (|a'\rangle\langle b'|c'\rangle\langle d'|)^\dagger\,. \tag{8.77}$$

Indeed,

$$(|a'\rangle\langle d'|)^\dagger = |a'd'|^\dagger = |d'a'| = |d'\rangle\langle a'|, \tag{8.78}$$

but the numerical coefficients are the same only if, generally,

$$(\lambda X)^\dagger = \lambda^* X^\dagger, \tag{8.79}$$

because $\langle c'|b'\rangle = \langle b'|c'\rangle^*$. Under the † operation, numbers are complex-conjugated. Thus † is a kind of extension of complex conjugation to operators.

Let's collect the rules we have learned:

$$(X+Y)^\dagger = X^\dagger + Y^\dagger, \tag{8.80a}$$
$$(XY)^\dagger = Y^\dagger X^\dagger, \tag{8.80b}$$
$$(\lambda X)^\dagger = \lambda^* X^\dagger. \tag{8.80c}$$

We can analyze $|a'b'|^\dagger = |b'a'|$ a bit further. In terms of vectors it reads

$$(|a'\rangle\langle b'|)^\dagger = |b'\rangle\langle a'|, \tag{8.81}$$

This means that † *reverses the order* of the vectors, and

$$|a'\rangle^\dagger = \langle a'|, \quad \langle b'|^\dagger = |b'\rangle. \tag{8.82}$$

Physically, this is obvious, since

- $\langle a'|$ represents putting the system into state a',
- $|a'\rangle$ reperesents taking the system out of state a'.

These two meanings are just interchanged by †.

We also note that

$$X^{\dagger\dagger} = X, \tag{8.83}$$

because reversing the convention of reading the order twice carries us back to where we started.

Finally, look at

$$|b'\rangle = \sum_{a'} |a'\rangle\langle a'|b'\rangle. \tag{8.84}$$

Take the adjoint of this:

$$|b'\rangle^\dagger = \langle b'| = \sum_{a'}\langle a'|b'\rangle^*\langle a'| = \sum_{a'}\langle b'|a'\rangle\langle a'|, \tag{8.85}$$

which is indeed correct.

Let's look at our spin-1/2 operators:

$$\sigma_x^\dagger = (|-\rangle\langle+| + |+\rangle\langle-|)^\dagger = |+\rangle\langle-| + |-\rangle\langle+| = \sigma_x, \tag{8.86a}$$

$$\sigma_y^\dagger = (i|-\rangle\langle+| - i|+\rangle\langle-|)^\dagger = -i|+\rangle\langle-| + i|-\rangle\langle+| = \sigma_y, \tag{8.86b}$$

$$\sigma_z^\dagger = (|+\rangle\langle+| - |-\rangle\langle-|)^\dagger = |+\rangle\langle+| - |-\rangle\langle-| = \sigma_z, \tag{8.86c}$$

$$1^\dagger = (|+\rangle\langle+| + |-\rangle\langle-|)^\dagger = |+\rangle\langle+| + |-\rangle\langle-| = 1 \tag{8.86d}$$

These operators are their own adjoints. We say that they are *self-adjoint* or *Hermitian*. The adjoint operation is often called Hermitian conjugation. To be self-adjoint is a requirement for any operator which represents a physical property:

$$A = \sum_{a'} a'|a'| = A^\dagger, \tag{8.87}$$

since a' is real and

$$|a'|^\dagger = (|a'\rangle\langle a'|)^\dagger = |a'\rangle\langle a'| = |a'|. \tag{8.88}$$

Could we have recognized the Hermitian property of the σ's from the corresponding matrices? Yes. To see how, consider a general operator expressed in terms of its matrix elements:

$$X = \sum_{a'a''} |a'\rangle\langle a'|X|a''\rangle\langle a''| = \sum_{a'a''} \langle a'|X|a''\rangle|a'a''|. \tag{8.89}$$

The adjoint is

$$X^\dagger = \sum_{a'a''} \langle a'|X|a''\rangle^* |a''a'| = \sum_{a''a'} \langle a''|X|a'\rangle^* |a'a''|, \tag{8.90}$$

where the last is obtained by relabelling the summation variables, $a' \leftrightarrow a''$. On the other hand, from the definition of the matrix elements of the operator $X^\dagger$,

$$X^\dagger = \sum_{a'a''} \langle a'|X^\dagger|a''\rangle|a'a''|, \tag{8.91}$$

so

$$\langle a'|X^\dagger|a''\rangle = \langle a''|X|a'\rangle^*. \tag{8.92}$$

The matrix of the adjoint operator $X^\dagger$ is obtained from the matrix of X by

1. interchanging rows and columns, and
2. taking the complex conjugate of each entry.

Thus, for matrices, the adjoint is the complex conjugate, transposed matrix. Transposition means that one interchanges rows and columns. A Hermitian matrix is equal to its own complex-conjugate, transpose. Thus

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_x^\dagger, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_y^\dagger, \tag{8.93a}$$

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_z^\dagger, \quad 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1^\dagger. \tag{8.93b}$$

With numbers $*$ acts like the adjoint:

$$\langle a''|X|a'\rangle^* = \langle a'|X^\dagger|a''\rangle, \quad \langle 1|2\rangle^* = \langle 2|1\rangle; \tag{8.94}$$

$*$ reverses the order of factors and takes the adjoint of each.

There is an intimate connection between the $\dagger$ operation and the U operators discussed previously. Recall that an operator is something which acts on a vector to give another vector. An operator is a rule which assigns to every vector another vector: it is a mapping. If the mapping *preserves the length* of *every* vector (analogous to a rotation of Euclidean vectors), it is called a *unitary* operator, customarily denoted U.

Let $\overline{\langle 1|}$ be the left vector which is the result of the mapping applied to $\langle 1|$:

$$\overline{\langle 1|} = \langle 1|U. \tag{8.95}$$

The corresponding right-vector statement is the adjoint of this:

$$\overline{|1\rangle} = U^\dagger|1\rangle. \tag{8.96}$$

If U preserves lengths, U is unitary:

$$\overline{\langle 1|1\rangle} = \langle 1|1\rangle, \tag{8.97}$$

or

$$\langle 1|UU^\dagger|1\rangle = \langle 1|1\rangle, \tag{8.98}$$

for all vectors $|1\rangle$. Clearly this last will be true if

$$UU^\dagger = 1. \tag{8.99}$$

In homework it is shown that this is a necessary result: A unitary operator is one for which

$$U^{-1} = U^\dagger. \tag{8.100}$$

We've seen this in particular examples:

$$U = e^{i\frac{\phi}{2}\sigma_z} = \cos\frac{\phi}{2} + i\sigma_z\sin\frac{\phi}{2}, \quad U^\dagger = \cos\frac{\phi}{2} - i\sigma_z\sin\frac{\phi}{2} = e^{-i\frac{\phi}{2}\sigma_z} = U^{-1}, \tag{8.101}$$

since $\sigma_z^\dagger = \sigma_z$.

This last example exhibits a general feature: Any unitary operator can be written as

$$U = e^{iH}, \quad H = H^\dagger, \quad U^\dagger = e^{-iH} = U^{-1}, \tag{8.102}$$

that is, the unitary operator U is written in terms of a Hermitian operator H. We'll not stop to prove this now.

Rotations not only preserve lengths, but all angles between vectors. (You can either rotate all vectors together, or the reference frame.) How is it with unitary transformations? If

$$\overline{\langle 1|} = \langle 1|U, \quad \overline{|2\rangle} = U^\dagger|2\rangle, \tag{8.103}$$

U preserves the scalar product:

$$\overline{\langle 1|2\rangle} = \langle 1|UU^\dagger|2\rangle = \langle 1|2\rangle \tag{8.104}$$

if $UU^\dagger = 1$.

We showed earlier that transformations of the type

$$\overline{X} = U^{-1}XU \tag{8.105}$$

preserve all algebraic relations

$$W = X + \alpha YZ \quad \text{implies} \quad \overline{W} = \bar{S} + \alpha\overline{YZ}. \tag{8.106}$$

If $U^{-1} = U^\dagger$ we have here the rule for the unitary transformation of operators. Is this consistent with the transformation law for vectors? That is, if $\langle 2|X = \langle 1|$ does

$$\overline{\langle 2|}\overline{X} = \langle 2|U(U^{-1}XU) = \langle 2|XU = \langle 1|U = \overline{\langle 1|}. \tag{8.107}$$

Similarly,

$$|1\rangle = X|2\rangle \quad \text{implies} \quad \overline{|1\rangle} = \overline{X}\overline{|2\rangle} \tag{8.108}$$

because

$$\overline{X}\overline{|2\rangle} = U^{-1}XUU^\dagger|2\rangle = U^{-1}X|2\rangle = U^\dagger|1\rangle = \overline{|1\rangle}. \tag{8.109}$$

What about adjoints of operators?

$$\overline{X} = U^{-1}XU = U^\dagger XU. \tag{8.110}$$

Take adjoint:

$$\overline{X}^\dagger = U^\dagger X^\dagger U = U^{-1}X^\dagger U = \overline{X^\dagger}, \tag{8.111}$$

since $U^{\dagger\dagger} = U$. The adjoint of the transformed operator is the transform of the adjoint. In particular, a Hermitian operator remains Hermitian under a unitary transformation: If

$$\overline{A} = U^\dagger AU, \quad \text{and} \quad A^\dagger = A, \quad \text{then} \quad \overline{A}^\dagger = \overline{A^\dagger} = \bar{A}. \tag{8.112}$$

Thus is physical property is maintained as a physical property under a unitary transformation.

All *numbers* (scalar products, eigenvalues, matrix elements) are maintained by unitary transformations:

$$\overline{\langle 1|2\rangle} = \langle 1|2\rangle, \tag{8.113a}$$
$$\overline{\langle 1|}\overline{X}\overline{|2\rangle} = \langle 1|UU^{-1}XUU^{-1}|2\rangle = \langle 1|X|2\rangle. \tag{8.113b}$$

A unitary transformation is a kind of rigid motion which does not change innter relations. A unitary transformation represents a change of description of the physical system; it represents a *freedom* in describing physical states and physical properties.

Physically, what are the freedom in describing systems? For example, we can translate or rotate our coordinate system, or go to a relatively moving coordinate frame. These freedoms, which have a awful lot to do with the physical world, have their description in terms of unitary transformations.

8.4 The Trace

The trace is an operation which maps operators into numbers; this is important, since numbers, which can be compared with experiment, are the goal of physical theory.

Consider the mapping

$$|a'\rangle\langle a''| \longrightarrow \langle a''|a'\rangle = \delta(a', a''), \tag{8.114}$$

which takes an operator, a measurement symbol, and assigns a number to it by simply reversing the order of the vectors. If we started with b vectors instead:

$$|b'\rangle\langle b''| \longrightarrow \langle b''|b'\rangle = \delta(b', b''). \tag{8.115}$$

Are these two statements consistent? After all, the b states can be constructed from the a states:

$$|b'\rangle = \sum_{a'} |a'\rangle\langle a'|b'\rangle, \tag{8.116a}$$

$$\langle b''| = \sum_{a''} \langle b''|a''\rangle\langle a''|, \tag{8.116b}$$

so

$$\begin{aligned}
|b'\rangle\langle b''| &= \sum_{a'} |a'\rangle\langle a'|b'\rangle \sum_{a''} \langle b''|a''\rangle\langle a''| \\
&\longrightarrow \sum_{a''} \langle b''|a''\rangle\langle a''| \sum_{a'} |a'\rangle\langle a'|b'\rangle \\
&= \sum_{a'a''} \langle b''|a''\rangle\langle a'|b'\rangle\langle a''|a'\rangle \\
&= \sum_{a'} \langle b''|a'\rangle\langle a'|b'\rangle = \langle b''|b'\rangle = \delta(b'', b').
\end{aligned} \tag{8.117}$$

Here's a final example of this mapping:

$$|a'\rangle\langle b'| \longrightarrow \langle b'|a'\rangle. \tag{8.118}$$

Check consistency of this by expressing b states in terms of a states:

$$\langle b'| = \sum_{a''} \langle b'|a''\rangle\langle a''| \tag{8.119}$$

so that

$$|a'\rangle\langle b'| = |a'\rangle \sum_{a''} \langle b'|a''\rangle\langle a''| \longrightarrow \sum_{a''} \langle b'|a''\rangle\langle a''|a'\rangle = \langle b'|a'\rangle, \tag{8.120}$$

since $\langle a''|a'\rangle = \delta(a', a'')$.

The operation is consistent; we have here a rule which maps operators into numbers, called the *trace*, abbreviated tr:

$$\operatorname{tr}|a'a''| = \delta(a', a''), \tag{8.121a}$$

$$\operatorname{tr}|a'| = 1, \tag{8.121b}$$

$$\operatorname{tr}|a'b'| = \langle b'|a'\rangle. \tag{8.121c}$$

More generally, if X is an arbitrary operator, expressed in terms of its matrix elements in the a description by

$$X = \sum_{a'a''} |a'\rangle\langle a'|X|a''\rangle\langle a''| = \sum_{a'a''} \langle a'|X|a''\rangle |a'a''|, \tag{8.122}$$

the trace of X is given by

$$\operatorname{tr} X = \sum_{a'a''} \langle a'|X|a''\rangle \delta(a', a''), \tag{8.123}$$

or

$$\operatorname{tr} X = \sum_{a'} \langle a'|X|a'\rangle. \tag{8.124}$$

That is, the trace is the diagonal sum, or the sum of the diagonal elements of the matrix of X. This coincides with the usual meaning of the trace of matrices.

Let's once again consider the spin-1/2 example, where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{8.125}$$

Then

$$\operatorname{tr} \sigma_x = \operatorname{tr} \sigma_y = \operatorname{tr} \sigma_y = 0. \tag{8.126}$$

These traces had to be the same, since the trace is the sum of the possible values a physical property can assume—see homework. On the other hand

$$\operatorname{tr} 1 = 2, \tag{8.127}$$

because there are two states.

In general, if there are n states, the unit operator is represented by the $n \times n$ unit matrix,

$$1 = \operatorname{diag}(1, 1, 1, \dots 1), \tag{8.128}$$

so

$$\operatorname{tr} 1 = n. \tag{8.129}$$

What about the trace of the product of two operators? Because

$$|a'a''||a'''a^{iv}| = \delta(a'', a''')|a'a^{iv}|, \tag{8.130}$$

$$\operatorname{tr}\left(|a'a''||a'''a^{iv}|\right) = \delta(a'', a''')\delta(a', a^{iv}). \tag{8.131}$$

In the other order

$$\begin{aligned}\mathrm{tr}\,\left(|a'''a^{iv}||a'a''|\right) &= \delta(a^{iv},a')\mathrm{tr}\,|a'''a''| = \delta(a^{iv},a')\delta(a'',a''') \\ &= \mathrm{tr}\,\left(|a'a''||a'''a^{iv}|\right). \end{aligned} \tag{8.132}$$

Because any operator is a linear combination of such measurement symbols, we have

$$\mathrm{tr}\,XY = \mathrm{tr}\,YX. \tag{8.133}$$

We can prove this in a different way, using matrix elements:

$$\mathrm{tr}\,XY = \sum_{a'}\langle a'|XY|a'\rangle = \sum_{a'a''}\langle a'|X|a''\rangle\langle a''|Y|a'\rangle, \tag{8.134a}$$

$$\mathrm{tr}\,YX = \sum_{a''}\langle a''|YX|a''\rangle = \sum_{a'a''}\langle a''|Y|a'\rangle\langle a'|X|a''\rangle = \mathrm{tr}\,XY. \tag{8.134b}$$

Check this with the properties of the σs:

$$\sigma_x\sigma_y = i\sigma_z, \quad \sigma_y\sigma_x = -i\sigma_z, \tag{8.135}$$

so

$$\mathrm{tr}\,i\sigma_z = i\,\mathrm{tr}\,\sigma_z = \mathrm{tr}\,\sigma_x\sigma_y = \mathrm{tr}\,\sigma_y\sigma_x = -i\,\mathrm{tr}\,\sigma_z, \tag{8.136}$$

which proves

$$\mathrm{tr}\,\sigma_z = 0 \tag{8.137}$$

without reference to the explicit construction of the σs. More generally,

$$\mathrm{tr}\,\sigma_k\sigma_l = \mathrm{tr}\,\sigma_l\sigma_k = \mathrm{tr}\,\frac{1}{2}\left(\sigma_k\sigma_l + \sigma_l\sigma_k\right). \tag{8.138}$$

But we know from Eq. (5.74) that

$$\frac{1}{2}\left(\sigma_k\sigma_l + \sigma_l\sigma_k\right) = \delta_{kl}1, \tag{8.139}$$

so we conclude

$$\mathrm{tr}\,\sigma_k\sigma_l = \delta_{kl}\mathrm{tr}\,1 = 2\delta_{kl}. \tag{8.140}$$

Another way of writing this is

$$\frac{1}{2}\mathrm{tr}\,(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{a}) = \mathbf{a}\cdot\mathbf{b}. \tag{8.141}$$

This follows from the identity (5.78),

$$(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{b}) = \mathbf{a}\cdot\mathbf{b} + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b}), \tag{8.142}$$

since $\mathrm{tr}\,\sigma_k = 0$.

Remember how probabilities were introduced in terms of measurement symbols:

$$|b'||a'||b'| = p(a',b')|b'|. \tag{8.143}$$

Take the trace of this:

$$\begin{aligned}\operatorname{tr}|b'||a'||b'| &= p(a',b')\operatorname{tr}|b'| = p(a',b') = \operatorname{tr}|b'||b'|a'| \\ &= \operatorname{tr}|b'||a'| = \operatorname{tr}|b'\rangle\langle b'|a'\rangle\langle a'| = \langle a'|b'\rangle\langle b'|a'\rangle, \end{aligned} \tag{8.144}$$

or

$$p(a',b') = |\langle a'|b'\rangle|^2, \tag{8.145}$$

as we know. But sometimes it is more convenient to use the intermediate form

$$p(a',b') = \operatorname{tr}|b'||a'|. \tag{8.146}$$

Again, return to spin 1/2. Let $\mathbf{e}_1$, $\mathbf{e}_2$ be the directions in which spin-measurements are made:

$$\begin{aligned}p(\sigma_1',\sigma_2') &= \operatorname{tr}|\sigma_1'||\sigma_2'| = \operatorname{tr}\frac{1+\sigma_1'\boldsymbol{\sigma}\cdot\mathbf{e}_1}{2}\frac{1+\sigma_2'\boldsymbol{\sigma}\cdot\mathbf{e}_2}{2} \\ &= \frac{1}{4}\operatorname{tr}\left(1+\sigma_1'\boldsymbol{\sigma}\cdot\mathbf{e}_1+\sigma_2'\boldsymbol{\sigma}\cdot\mathbf{e}_2+\sigma_1'\sigma_2'\boldsymbol{\sigma}\cdot\mathbf{e}_1\boldsymbol{\sigma}\cdot\mathbf{e}_2\right) \\ &= \frac{1}{4}\left(2+0+0+\sigma_1'\sigma_2'2\mathbf{e}_1\cdot\mathbf{e}_2\right) \\ &= \frac{1}{2}\left(1+\sigma_1'\sigma_2'\mathbf{e}_1\cdot\mathbf{e}_2\right) = \frac{1}{2}(1+\sigma_1'\sigma_2'\cos\Theta), \end{aligned} \tag{8.147}$$

where Θ is the angle between $\mathbf{e}_1$ and $\mathbf{e}_2$. Thus we recover

$$p(\pm1,\pm2) = \cos^2\frac{\Theta}{2}, \quad p(\mp1,\pm2) = \sin^2\frac{\Theta}{2}, \tag{8.148}$$

as we've seen many times.