Chapter 6

Probabilities and vectors

Now we must learn how to extract probabilities from our measurement symbols. Thus, suppose we consider two successive Stern-Gerlach experiments, one of which measures (selects) a particular vlaue of $\sigma_{z'}$, $\sigma''_{z'}$, and a second which selects a particular value of σ_z, σ'_z. The two directions, z, z', are in general not the same. We want to learn how to calculate $p(\sigma'_z, \sigma''_{z'})$, the probability of finding $\sigma_z = \sigma'_z$ given that the first measurement obtained $\sigma_{z'} = \sigma''_{z'}$.

In general, suppose we have two different properties A and B. We first select systems with $B = b''$. What is the probability of subsequently measuring $A = a'$,

$$p(a', b'')?$$

We have in mind here that a measurement of A *or* a measurement of B describes the system fully. If we know the value of either A or B we know all we can about the system. We cannot measure A and B simultaneously, measurement of A destroys what is known abouty B. This is precisely the situation with σ_z, $\sigma_{z'}$.

The sequence of selective measurement described above is represented by

$$|a'||b''|. \tag{6.1}$$

The second measuremnt destroys at least some of the information determined by the first measurement. (Thus, measurement of J_z destroys, in an uncontrollable way, information about J_x.) To see what happens as a result of the A measurement, we remeasure B:

$$|b''||a'||b''|. \tag{6.2}$$

This measures the effect of the disturbance produced by the intermediate A measurment. The net effect is a selection of b'':

$$|b''||a'||b''| = \text{number}|b''|, \tag{6.3}$$

since the overall effect is a selection of b'' and an emission of b''. What is the number here? If $A = B$, $b'' = a''$,

$$|a''||a'||a''| = \delta(a', a'')|a''|, \tag{6.4}$$

that is, if $a'' = a'$, we are just repeating the same selective measurement. Then, the number is 1, which expresses the *certainty* of the second measurement giving $A = a'$. If $a' \neq a''$, the second measurement rejects all that comes from the first; the second measurement will *certainly not* find any systems with the property $A = a'$.

This example makes us suspect that

$$|b''||a'||b''| = p(a', b'')|b''|, \tag{6.5}$$

where $p(a', b'')$ is the probability of finding a' given b'', In the above example,

$$p(a', a'') = \delta(a', a''). \tag{6.6}$$

We will now verify that this is a consistent hypothesis.

Suppose the intermediate measurement is a less selective one, in which either $A = a'$ or $A = a''$ is selected without discrimination:

$$|b''|(|a'| + |a''|)|b''| = [p(a', b'') + p(a'', b'')]\,|b''|, \tag{6.7}$$

where the intermediate measurement corrsponds to a measurement of $(A - a')(A - a'')$. This make sense, because $p(a', b'')$ is the fraction of atoms with $A = a'$, $p(a'', b'')$ is the fraction of atoms with $a = a''$, while $p(a', b'') + p(a'', b'')$ is the fraction of atoms with either $A = a'$ or $A = a''$. Fractions, probabilities, add, as they must.

Suppose the intermediate measurement is a completely non-selective one,

$$1 = \sum_{a'} |a'|, \tag{6.8}$$

which selects all possible outcomes. Then

$$|b''| \sum_{a'} |a'||b''| = \sum_{a'} p(a', b'')|b''| = |b''|, \tag{6.9}$$

so

$$\sum_{a'} p(a', b'') = 1. \tag{6.10}$$

The sum of the probabilites for all possible outcomes is unity.

A further property that probabilities must satisfy is that they be non-negative. We will prove this later after we develop some more machinery. For now, we will simply accept the above identification.

6.1 Example

Suppose

$$b'' = \sigma''_{z'} = +1, \quad a' = \sigma'_z = +1. \tag{6.11}$$

Then is

$$|\sigma''_{z'} = +1|\sigma'_z = +1||\sigma''_{z'} = +1| = p(+z, +z')|\sigma''_{z'} = +1|. \tag{6.12}$$

Here $p(+z,+z')$ is the probability of finding $\sigma'_z = +1$ if previously the measurment determined $\sigma''_{z'} = +1$. In terms of σ's

$$\frac{1+\sigma_{z'}}{2}\frac{1+\sigma_z}{2}\frac{1+\sigma_{z'}}{2} = p(+z,+z')\frac{1+\sigma_{z'}}{2}. \tag{6.13}$$

But we know

$$\sigma_z = \sigma_{x'}\sin\theta\cos\phi + \sigma_{y'}\sin\theta\sin\phi + \sigma_{z'}\cos\theta, \tag{6.14}$$

in terms of the polar angles of $\hat{\mathbf{z}}$ relative to the x', y', z' coordinate system. The x' and y' components in fact don't matter here:

$$\frac{1+\sigma_{z'}}{2}\sigma_{x'}\frac{1+\sigma_{z'}}{2} = \sigma_{x'}\frac{1-\sigma_{z'}}{2}\frac{1+\sigma_{z'}}{2} = 0, \tag{6.15}$$

since $\sigma_{x'}$ anticommutes with $\sigma_{z'}$ amd $\sigma_{z'}^2 = 1$. (The vanishing of the above uses

$$|\sigma_{z'} = -1||\sigma_{z'} = +1| = 0.) \tag{6.16}$$

Thus

$$\begin{aligned}\frac{1+\sigma_{z'}}{2}\frac{1+\sigma_z}{2}\frac{1+\sigma_{z'}}{2} &= \frac{1+\sigma_{z'}\cos\theta}{2}\frac{1+\sigma_{z'}}{2}\frac{1+\sigma_{z'}}{2} \\ &= \frac{1+\sigma_{z'}\cos\theta}{2}|\sigma'_{z'} = +1| \\ &= \frac{1+\cos\theta}{2}|\sigma'_{z'} = +1|,\end{aligned} \tag{6.17}$$

from which we conclude

$$p(+z,+z') = \frac{1+\cos\theta}{2} = \cos^2\frac{\theta}{2}, \tag{6.18}$$

as we determined previously. The fact that

$$P(-z,+z') = \sin^2\frac{\theta}{2} \tag{6.19}$$

now follows from

$$\sum_{\sigma'_z} p(\sigma'_z, \sigma''_{z'}) = 1. \tag{6.20}$$

The machinery works.

6.2 State vectors

Now we proceed with a further development of our mathematical transcription of the physical acts of measurement. Recall that we discussed measurements which changed the value of a physical property, represented by the symbol $|a'a''|$: which selects $A = a'$, and emits $A = a''$. A Stern-Gerlach representation of this was sketched in Fig. 2.6. There, for example, the symbol $\otimes$ represented a rotation of the spin of the atom about the x axis through $180°$, for example, by precession about a longitudinal magnetic field. But this picture is misleading, because we cannot measure what the spin is doing internally. Such an internal measurement would destory what we are trying to do here. You would be making a different measurement. We can only say what goes in and what goes out.

What do we mean by $|a'a''|$? The measurement takes atoms from the state a' and puts them into the state a''. We can think of this in two stages, as long as we don't try to state how the transition occurs:

1. Atoms are removed from a'.
2. Atoms are put into a''.

The algebraic counterpart of this two-stage process is written as

$$|a'a''| = |a'\rangle\langle a''|, \tag{6.21}$$

where $|a'\rangle$ represents removal of an atom from a', $\langle a''|$ represents placing atoms in a''.

We must insist that we cannot analyze the transition between the states a', a'' in time. Is this factorization consistent with everything that went before? Remember

$$|a'a''||a'''a^{iv}| = \delta(a'', a''')|a'a^{iv}|. \tag{6.22}$$

In the new notation, this becomes

$$|a'\rangle\langle a''||a'''\rangle\langle a^{iv}| = |a'\rangle\langle a''|a'''\rangle\langle a^{iv}| = \delta(a'', a''')|a'\rangle\langle a^{iv}|, \tag{6.23}$$

where we have simplified the notation by coalescing the common vertical line. The above equality is true provided

$$\langle a''|a'''\rangle = \delta(a'', a'''), \tag{6.24}$$

which is a number, describing an internal rearrangement. We interpret this number $\langle a''|a'''\rangle$ as follows: Start with no atoms; put atoms out in state a''; absorb atoms is state a'''; put out no atoms. If $a'' \neq a'''$ there is nothing to absorb, so we get 0.

We now see that every physical state has two "vectors" associated with it:

$$|a'\rangle, \quad \langle a''|. \tag{6.25}$$

These are unit, orthogonal vectors

$$\langle a'|a'\rangle = 1, \quad \langle a'|a''\rangle = 0, \quad a' \neq a''. \tag{6.26}$$

Unlike with ordinary vectors, we don't multiply vectors by themselves, but one kind of vector with a different kind of vector. We have a *complex* geometry;

$\langle a'|$ is some sort of *complex conjugate* of $|a'\rangle$. The geometry is n-dimensional if there are n values of a'.

Think about 3-dimensional Euclidean geometry. We have three unit orthogonal vectors, say $\mathbf{e}_k$, $k = 1, 2, 3$:

$$\mathbf{e}_k \cdot \mathbf{e}_l = \delta_{kl}. \tag{6.27}$$

We can write an arbitrary vector in terms of these basis vectors:

$$\mathbf{V} = \sum_{k=1}^{3} \mathbf{e}_k v_k, \tag{6.28}$$

where the components, v_k are the projections of $\mathbf{V}$ on the axes defined by $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$v_k = \mathbf{e}_k \cdot \mathbf{V}. \tag{6.29}$$

Putting these two expression together gives

$$\mathbf{V} = \sum_k \mathbf{e}_k \mathbf{e}_k \cdot \mathbf{V} = \left(\sum_k \mathbf{e}_k \mathbf{e}_k \right) \cdot \mathbf{V} = \mathbf{1} \cdot \mathbf{V}, \tag{6.30}$$

where we see the construction of $\mathbf{1}$, the unit dyadic (unit symbol) in terms of the $\mathbf{e}$'s:

$$\mathbf{1} = \sum_k \mathbf{e}_k \mathbf{e}_k. \tag{6.31}$$

This must hold in any coordinate system, that is, for any set of unit, orthogonal vectors. Since the above property of $\mathbf{1}$ holds for any vector V, it implies $\mathbf{1} \cdot \mathbf{e}_k = \mathbf{e}_k$ in particular.

This is all very analogous to what we have in our algebra of measurement. There the unit symbol is

$$1 = \sum_{a'} |a'| = \sum_{a'} |a'a'| = \sum_{a'} |a'\rangle\langle a'|. \tag{6.32}$$

Now, however, in the sum, we have the two different kinds of vectors sitting side by side.

(If, instead, we had the vectors in the other order, it would mean

$$\sum_{a'} \langle a'|a'\rangle = \sum_{a'} 1 = n, \tag{6.33}$$

the number of states.)

From the physical meaning of 1 as a totally non-selective measurement, we know

$$1|a'a''| = |a'a''|1 = |a'a''|. \tag{6.34}$$

That is,

$$1|a'\rangle\langle a''| = |a'\rangle\langle a''|, \quad |a'\rangle\langle a''|1 = |a'\rangle\langle a''|. \tag{6.35}$$

In the first case, $\langle a''|$ is inert—nothing happens to it; while in the second, $|a'\rangle$ is inert. Thus we conclude

$$1|a'\rangle = |a'\rangle, \quad \langle a''|1 = \langle a''|. \tag{6.36}$$

Let's see how this works out:

$$1|a'\rangle = \left(\sum_{a''} |a''\rangle\langle a''|\right)|a'\rangle = \sum_{a''} |a''\rangle\langle a''|a'\rangle = \sum_{a''} |a''\rangle\delta(a'', a') = |a'\rangle, \tag{6.37a}$$

$$\langle a''|1 = \langle a''|\sum_{a'} |a'\rangle\langle a'| = \sum_{a'}\langle a''|a'\rangle\langle a'| = \sum_{a'} \delta(a'', a')\langle a'| = \langle a''|. \tag{6.37b}$$

These are just as expected.

Nomenclature: We call $|a'\rangle$ a *right vector* and $\langle a''|$ a *left vector*, since in

$$\langle a''|a'\rangle \tag{6.38}$$

$\langle a''|$ is on the left, $|a''\rangle$ is on the right. (Dirac called these *bra* and *ket* vectors, since together they form a *bra(c)ket.*) Each physical state has both a left and a right vector associated with it.

Back to three-dimensional Euclidean space. We are utterly free to choose a different coordinate system, defined by the basis vectors

$$\mathbf{e}_{k'}, \quad k' = 1, 2, 3. \tag{6.39}$$

Equally well with these,

$$\sum_{k'} \mathbf{e}_{k'}\mathbf{e}_{k'} = \mathbf{1} \tag{6.40}$$

is the unit dyadic. How are the unit vectors of one coordinate system related to the unit vectors of another?

$$\mathbf{e}_k = \mathbf{1}\cdot\mathbf{e}_k = \left(\sum_{l'} \mathbf{e}_{l'}\mathbf{e}_{l'}\right)\cdot\mathbf{e}_k = \sum_{l'} \mathbf{e}_{l'}(\mathbf{e}_{l'}\cdot\mathbf{e}_k). \tag{6.41}$$

We call $(\mathbf{e}_{l'}\cdot\mathbf{e}_k)$ is a "direction cosine;" it is the cosine of the angle between $\mathbf{e}_k$ and $\mathbf{e}_{l'}$. Or we could do this the other way:

$$\mathbf{e}_{k'} = \mathbf{1}\cdot\mathbf{e}_{k'} = \left(\sum_{l} \mathbf{e}_{l}\mathbf{e}_{l}\right)\cdot\mathbf{e}_{k'} = \sum_{l} \mathbf{e}_{l}(\mathbf{e}_{l}\cdot\mathbf{e}_{k'}); \tag{6.42}$$

$(\mathbf{e}_l\cdot\mathbf{e}_{k'})$ also being a direction cosine, the cosine of the angle between $\mathbf{e}_l$ and $\mathbf{e}_{k'}$. Here the direction cosines which express the old vectors in terms of the new vectors, and the new vectors in terms of the old vectors, are exactly the same.

How is it with our measurement algebra? Here, changing the coordinate system corresponds to measuring different physical quantities—such as components of angular momentum in one, or another, direction. Supose we have

- physical quantities: $A \qquad B \quad$ (e.g., J_z, $J_{z'}$)
- typical results of measurements: $a' \quad b'$
- corresponding states: $|a'\rangle$, $\langle a'| \quad |b'\rangle$, $\langle b'|$.

If we measure property A and get all possible states (no discrimination), we have the unit symbol, representing no measurement at all. Similarly with B. So

$$1 = \sum_{a'} |a'\rangle\langle a'| = \sum_{b'} |b'\rangle\langle b'|. \tag{6.43}$$

The states are represented by orthogonal unit vectors:

$$\langle a'|a''\rangle = \delta(a', a''), \quad \langle b'|b'\rangle = \delta(b', b''). \tag{6.44}$$

$\{|a'\rangle\}$, $\{|b'\rangle\}$ are two different orthonormal sets of vectors. We now transform from one of these orthonormal sets to the other:

$$\langle a'| = \langle a'| \sum_{b''} |b''\rangle\langle b''| = \sum_{b''} \langle a'|b''\rangle\langle b''| = \langle a'|. \tag{6.45}$$

This expresses how a $\langle a'|$ vector is written in terms of $\langle b''|$ vectors. The expansion coefficients are scalar products of a and b vectors—these numbers express how the old coordinate system is related to the new coordinate system. In other words, $\langle a'|b''\rangle$ is a kind of direction cosine.

How does the transformation work with right vectors?

$$|a'\rangle = 1|a'\rangle = \sum_{b''} |b''\rangle\langle b''|a'\rangle. \tag{6.46}$$

This expresses right $|a'\rangle$ vectors in terms of right $|b''\rangle$ vectors, in terms of "direction cosines" $\langle b''|a'\rangle$. But unlike in Euclidean geometry, here there are two kinds of direction cosines,

$$\langle b''|a'\rangle \neq \langle a'|b''\rangle. \tag{6.47}$$

The second kind also appears in

$$\langle b''|1 = \langle b''| \sum_{a'} |a'\rangle\langle a'| = \sum_{a'} \langle b''|a'\rangle\langle a'|, \tag{6.48}$$

the expansion of left b vectors in terms of left a vectors.

To find the relation between these two kinds of direction cosines, we consider the probability of finding a particular value of A after selecting a state with a definite value of B:

$$|b''||a'||b''| = p(a', b'')|b''|, \tag{6.49}$$

where $p(a', b'')$ is the fraction of atoms with $A = a'$ if the atoms were previously selected to have $B = b''$. Analyze this into six stages:

$$|b''\rangle\langle b''|a'\rangle\langle a'|b''\rangle\langle b''| = p(a', b'')|b''\rangle\langle b''|. \tag{6.50}$$

Remove the common left and right b vectors, and we have

$$\langle b''|a'\rangle\langle a'|b''\rangle = p(a', b''). \tag{6.51}$$

By simply interchanging the letters, $b'' \leftrightarrow a'$,

$$\langle a'|b''\rangle\langle b''|a'\rangle = p(b'', a'), \tag{6.52}$$

so since the right side of this equation is simply the product of the same two *numbers*, we conclude that the probabilities are symmetric,

$$p(a', b'') = p(b'', a'), \tag{6.53}$$

as we have already seen in spin-1/2, 1 examples. The probabilities are constructed as the product of both kinds of direction cosines.

Now impose the physical requirement that probabilities be *real, non-negative* numbers. This will automatcally be true (and it is essentially impossible to think of any other possibility that is not reducible to this) if

$$\langle a'|b''\rangle = \langle b''|a'\rangle^*, \tag{6.54}$$

for then

$$p(a', b'') = |\langle a'|b''\rangle|^2 \geq 0. \tag{6.55}$$

(We have already seen in examples that these "amplitudes" cannot be real.) That is, the two direction cosines are complex conjugates of each other. This is the simplest possibility beyond Euclidean geometry, where the direction cosines are equal. Note that

$$\langle a'| = \sum_{b''}\langle a'|b''\rangle\langle b''|, \quad |a'\rangle = \sum_{b''}|b''\rangle\langle b''|a'\rangle = \sum_{b''}|b''\rangle\langle a'|b''\rangle^*. \tag{6.56}$$

This shows that left and right vectors are in some sense complex conjugates of each other.

Now multiply the above expansion for a left vector by a right vector,

$$\left(\langle a'| = \sum_{b'}\langle a'|b'\rangle\langle b'|\right)|a''\rangle, \tag{6.57}$$

or

$$\langle a'|a''\rangle = \delta(a', a'') = \sum_{b'}\langle a'|b'\rangle\langle b'|a''\rangle. \tag{6.58}$$

This means that the sums of the absolute squares of the "direction cosines" is unity (geometry)

$$1 = \sum_{b'}|\langle a'|b'\rangle|^2 = \sum_{b'}p(a', b') = \sum_{a'}p(a', b'), \tag{6.59}$$

where the last follows from the symmetry property (6.53). This, physically, is the fact, built in, that the sum of the probabilities for all outcomes is 1—*some* outcome is certain.

Similarly, we have

$$\langle b'|b''\rangle = \sum_{a'} \langle b'|a'\rangle\langle a'|b''\rangle = \delta(b', b''). \tag{6.60}$$

This last equality is a statement that the b vectors are orthonormal, given that the a vectors are:

$$\langle b'| = \sum_{a'} \langle b'|a'\rangle\langle a'|, \quad |b''\rangle = \sum_{a''} |a''\rangle\langle a''|b''\rangle, \tag{6.61}$$

so the inner product of these vectors is

$$\langle b'|b''\rangle = \sum_{a'a''} \langle b'|a'\rangle\langle a'|a''\rangle\langle a''|b''\rangle = \sum_{a'} \langle b'|a'\rangle\langle a'|b''\rangle = \delta(b', b''), \tag{6.62}$$

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

Now note that

$$\sum_{b'}\sum_{a'} \langle b'|a'\rangle\langle a'|b'\rangle = \sum_{b'} 1 = n_b, \tag{6.63a}$$

$$\sum_{a'}\sum_{b'} \langle a'|b'\rangle\langle b'|a'\rangle = \sum_{a'} 1 = n_a, \tag{6.63b}$$

where n_a is the number of a states, and n_b is the number of b states. The two sums are equal, so we conclude that $n_a = n_b = n$, the number of states of the system, which is independent of which property is used to define them.

Following Dirac, we call the set of n^2 numbers $\langle a'|b'\rangle$ the *transformation function*. In terms of these numbers, we transform from one coordinate system to another. We call $\langle a'|b'\rangle$ the ab transformation function, and $\langle b'|a'\rangle$ the ba transformation function, which is the complex conjugate of the ab transformation function.

The ba transformation function undoes, is the inverse of, the ab transformation function. For if we do the two transformation in succession,

$$|a''\rangle = \sum_{a'b'} |a'\rangle\langle a'|b'\rangle\langle b'|a''\rangle = \sum_{a'} |a'\rangle \sum_{b'} \langle a'|b'\rangle\langle b'|a''\rangle = \sum_{a'} |a'\rangle\delta(a', a''). \tag{6.64}$$

The net transformation is no transformation at all—the identity transformation.

Suppose we now go from a description in terms of A states to one in terms of B states and finally to one in C states:

$$a \to b: \quad |b'\rangle = \sum_{a'} |a'\rangle\langle a'|b'\rangle, \tag{6.65a}$$

$$b \to c: \quad |c'\rangle = \sum_{b'} |b'\rangle\langle b'|c'\rangle, \tag{6.65b}$$

so going from $a \to b \to c$,

$$|c'\rangle = \sum_{a'b'} |a'\rangle\langle a'|b'\rangle\langle b'|c'\rangle = \sum_{a'} |a'\rangle\langle a'|c'\rangle, \tag{6.66}$$

where in the latter expresses the direct transformation $a \to c$. Thus

$$\langle a'|c'\rangle = \sum_{b'} \langle a'|b'\rangle\langle b'|c'\rangle \tag{6.67}$$

is the composition property of transformation functions. It says how to construct the ac transformation function in terms of the ab and bc transformation functions. Note that there is a slick method of getting this result:

$$\langle a'|c'\rangle = \langle a'|1|c'\rangle = \langle a'|\sum_{b'}|b'\rangle\langle b'||c'\rangle = \sum_{b'}\langle a'|b'\rangle\langle b'|c'\rangle. \tag{6.68}$$

Finally, note that this composition property is compatible with

$$\langle a'|b'\rangle = \langle b'|a'\rangle^* \tag{6.69}$$

because taking the complex conjungate of Eq. (6.68) give

$$\langle a'|c'\rangle^* = \sum_{b'}\langle a'|b'\rangle^*\langle b'|c'\rangle^*, \tag{6.70}$$

which is the same as

$$\langle c'|a'\rangle = \sum_{b'}\langle c'|b'\rangle\langle b'|a'\rangle, \tag{6.71}$$

because the two factors in the summand are simply interchanged.