Chapter 5

Construction of Quantum Mechanics

We are now going to develop the mathematical framework of quantum mechanics, which is tied fundamentally to the process of measurement. It is a symbolic representation of experiment. We generalize from what is done in the Stern-Gerlach experiment, where a property μ_z or J_z is measured. In general, we'll say we measure a physical property A, and the results of the measurement are the possible values of A, which are real numbers:

$$\text{possible values of A:} \quad a_1, a_2, \dots, a_n, \tag{5.1}$$

and we'll denote a typical value by a' or a''. A measurement is also a selection. In the Stern-Gerlach experiment, we select a particular value of J_z, stopping all other sub-beams with different J_z values. In general, the experiment has the schematic form shown in Fig. 5.1. Of course, atoms have more than one property—we here disregard all but one, A. The beam coming out of the measurement apparatus is said to be in a definite state in which $A = a'$.

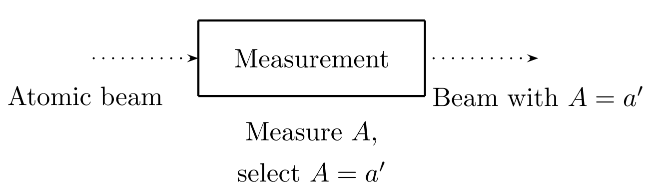

Figure 5.1: A general selective measurement. An arbitrary beam of atoms enters the apparatus, which measures the property A, and selects those atoms that have the value $A = a'$. The atoms that leave the apparatus all have the property $A = a'$.

To represent this measurement, we introduce the symbol, the *measurement symbol*:

$$|a'|. \tag{5.2}$$

(This is not to be confused with an absolute value sign.) It represents a selective measurement in which the property A is measured, and only those atoms which have $A = a'$ are selected.

If we follow one Stern-Gerlach experiment by an identical one, nothing happens: all atoms entering the first apparatus are selected by the second, as seen in Fig. 2.5. The second measurment simply verifies the first. Symbolically, we express this as

$$|a'|\,|a'| = |a'|. \tag{5.3}$$

On the other hand, if we measure a different state, we get nothing, The general symbolic statement of this is

$$|a'|\,|a''| = 0, \quad \text{if} \quad a' \neq a'', \tag{5.4}$$

where 0 is the symbol of a measurement that rejects everything. What the first selects, the second rejects. (It makes no difference if these equations are read from right to left or from left to right.) The symbol 0 has the following properties,

$$\begin{aligned} |a'|\,0 &= 0, \\ 0\,|a'| &= 0, \\ 0\,0 &= 0. \end{aligned} \tag{5.5}$$

The first two equations say that if we attempt to measure a property, before or after rejecting everything, you get nothing. We are beginning to see an algebra, in which the multiplication of symbols represents performing one experiment after another.

Now, what do we mean by the addition of measurement symbols,

$$|a'| + |a''| = ? \tag{5.6}$$

It represents a less selective measurement, in which the selected atoms have *either* property $A = a'$ or $A = a''$ without discrimination. We *do not* mean that you measure a', a'' separately, and put these selected "beams" back together. No separation of atoms with property $A = a'$ or $A = a''$ is made. Intervention by measurment is a dramatic event. Here, we do not distinguish a' from a'' in any way.

By the physical meaning of addition, the order does not matter:

$$|a'| + |a''| = |a''| + |a'|. \tag{5.7}$$

Similarly, we could perform an even less selective measurement in which $A = a', a''$, or a''' is selected without discrimination. This is represented by the symbol

$$|a'| + |a''| + |a'''|, \tag{5.8}$$

where the terms can be written in any order. We can keep going in this manner until we select atoms which have any value of the property A, without discrimination. Then we reject no atoms. The symbolic transcription of this is

$$|a_1| + |a_2| + |a_3| + \ldots + |a_n| = \sum_{a'} |a'| = 1, \tag{5.9}$$

where the summation sign means that we sum over all possible values of a' from a_1 through a_n, and where 1 is a symbol for a measurement that selects all systems without discrimination—that is, no measurement at all, since nothing is done to the "beam." The properties of 1 are evident, since it corresponds to letting everything through:

$$\begin{aligned} |a'|\,1 &= |a'|, \\ 1\,|a'| &= |a'|, \\ 1\,1 &= 1, \\ 1\,0 &= 0, \\ 0\,1 &= 0, \end{aligned} \tag{5.10}$$

1 has the algebraic properties of unity.

Next we ask, does the distrbutive property hold in this new algebra? For example, does the following equation hold:

$$\left(\sum_{a'} |a'|\right) |a''| = \sum_{a'} (|a'||a''|)? \tag{5.11}$$

The left side of this equation is

$$1\,|a''| = |a''|. \tag{5.12}$$

On the right side, all terms are 0 except $|a''||a''| = |a''|$, so the right side is

$$|a''| + 0 + 0 + \ldots + 0 = |a''|, \tag{5.13}$$

because $|a''| + 0$ means we either we select $A = a''$ *or* reject everything, which means we simply select $A = a''$. This indicates that the distributive property holds. (The general distributive property is proved in homework.)

Now we want to introduce a symbol for the physical quantity A itself. Since $|a'|$ represents a "filtration" of the beam, filtering out only those atoms in which $A = a'$, we let the symbol for the property A, also called A, satisfy

$$A|a'| = a'|a'|; \tag{5.14}$$

on the left side we have the symbol for the property A, and on the right we have the numerical value A assumes in that state, a'. This means that if we first select atoms with property $A = a'$ and then measure A, we will of course get the value a'. We can read these symbols either way, so we also have

$$|a'|A = |a'|a'. \tag{5.15}$$

From this, we can write A explicitly in terms of the measurement symbols $|a_1|$, $\dots$, $|a_n|$:

$$\sum_{a'} A|a'| = \sum_{a'} a'|a'|, \tag{5.16}$$

where on the left we have

$$A\sum_{a'} |a'| = A\,1 = A, \tag{5.17}$$

since multiplication is distributive. This exhibits A:

$$A = \sum_{a'} a'|a'|. \tag{5.18}$$

Indeed, this is consistent,

$$A|a''| = \sum_{a'} a'|a'||a''| = a''|a''|, \tag{5.19}$$

since every term in the sum is 0, except for $a''|a''||a''| = a''|a''|$. Note that our definition of A means that

$$0|a'| = 0, \tag{5.20}$$

where on the left 0 represents the number zero, while on the right appears the symbol 0. This is because $A - a'$ is also a physical quantity, and if A has the value a', $A - a'$ has the value 0 (we are just shifting the origin):

$$(A - a')|a'| = 0|a'| = 0; \tag{5.21}$$

this is just a rewriting of $A|a'| = a'|a'|$.

Now we have a small check of consistency. If a typical value of A is a', the corresponding value of A^2 is $(a')^2$. In terms of measurement symbols, this means

$$A^2 = \sum_{a'} (a')^2|a'|, \tag{5.22}$$

since a state in which A has the value a' is a state in which A^2 has the value $(a')^2$. Is this really the square of A?

$$A^2 = A\sum_{a'} a'|a'| = \sum_{a'} a'A|a'| = \sum_{a'} (a')^2|a'|. \tag{5.23}$$

Obviously, we could keep on taking powers of A. In general define a function of A, $f(A)$, by

$$f(A) = \sum_{a'} f(a')|a'|. \tag{5.24}$$

This means that $f(A)$ has the value $f(a')$ when we select $A = a'$,

$$f(A)|a'| = f(a')|a'|. \tag{5.25}$$

Here are some examples.

- Suppose for all a', $f(a') = 1$, then
$$f(A) = \sum_{a'} 1|a'| = \sum_{a'} |a'| = 1, \tag{5.26}$$
the function is just the unit symbol.

- If $f(a') = 0$ for all a',
$$f(A) = \sum_{a'} 0|a'| = 0, \tag{5.27}$$
the function is the zero zymbol.

- If $f(a') = 1$ for a particular a', but zero for all others, $f(a'') = 0$, $a'' \neq a'$, then
$$f(A) = |a'|. \tag{5.28}$$
This means, that the measurement symbol is a function of A. In fact, in the homework you will prove that
$$|a'| = \prod_{a'' \neq a'} \frac{(A - a'')}{(a' - a'')}. \tag{5.29}$$

Now let's return to our doublet, spin-1/2, two-level system. Let us write the z-component of angular momentum as
$$J_z = \frac{1}{2}\hbar\sigma_z, \quad \sigma_z' = \pm 1, \tag{5.30}$$
where we now regard σ_z as a symbol, while following the above notation the possible values of σ_z are denoted by a prime. The symbol σ_z is written in terms of measurment symbols as
$$\sigma_z = +1| + 1| + (-1)| - 1| = | + | \; - \; | - |. \tag{5.31}$$
Where we have simplified the notation to label the state by the sign of σ_z'. On the other hand,
$$1 = | + | \; + \; | - |. \tag{5.32}$$
Add and subtract these:
$$| + | = \frac{1 + \sigma_z}{2}, \quad | - | = \frac{1 - \sigma_z}{2}. \tag{5.33}$$
Let us check that the required properties hold:
$$| + || - | = \frac{1 - \sigma_z^2}{4} = 0, \tag{5.34}$$
since σ_z^2 always has the value 1, while
$$| + || + | = \left(\frac{1 + \sigma_z}{2}\right)^2 = \frac{1 + 2\sigma_z + \sigma_z^2}{4} = \frac{1 + \sigma_z}{2} = | + |, \tag{5.35a}$$
$$| - || - | = \left(\frac{1 - \sigma_z}{2}\right)^2 = \frac{1 - 2\sigma_z + \sigma_z^2}{4} = \frac{1 - \sigma_z}{2} = | - |. \tag{5.35b}$$

5.1 General measurement symbols

The algebra developed to this point is too special. Let's consider measurements in which atoms are selected in one state, and emitted in another state, that is, where there is a change of state. Fig. 2.6 shows an example. There, a Stern-Gerlach measurement which selects atom in a state $m = +1/2$ is followed by an apparatus which changes all those selected atoms into the $m = -1/2$ state. A subsequent Stern-Gerlach apparatus verifies that all the atoms in the selected beam have $m = -1/2$. Let's convince ourselves that such a change in state is possible. We can imagine at least two ways of changing the $m = +1/2$ state to the $m = -1/2$ state.

1. Apply a magnetic field perpendicular to the z-axis, the direction of the magnetic moment selected by the first Stern-Gerlach apparatus. This will cause the magnetic moment, or the spin, to precess around the direction of the field, and that precession can be arranged so the precession is through π or 180°. Then $J_z = +\hbar/2$ gets converted to $J_z = -\hbar/2$.

2. Apply a magnetic field parallel to the z-axis, that is, parallel to the direction of the selected magnetic moment. Because the energy of the magnetic dipole in the magnetic field is $E = -\boldsymbol{\mu} \cdot \mathbf{H}$, the $m = -1/2$ and $m = +1/2$ states have different energies, and if $\gamma > 0$, the former has higher energy. Now send in an appropriate electromagnetic wave to cause a transition from $m = +1/2$ to $m = -1/2$, which could be done by feeding in electromagnetic energy at resonance,

$$E_{\text{em wave}} = E_{m=-1/2} - E_{m=+1/2}. \tag{5.36}$$

   This technique is used in the standard of time, the cesium fountain clock.[1]

We now generalize, and consider a measurement of a property A. We select atoms with $A = a'$ but emit them with $A = a''$. The measurement symbol for this is

$$|a'a''|. \tag{5.37}$$

(This symbol can be read in either order.) This generalizes what we had before. When there is no change in state,

$$|a'a'| = |a'|. \tag{5.38}$$

Suppose we follow one such measurement by another such,

$$|a'a''||a''a'''| = |a'a'''|, \tag{5.39}$$

[1] Since 1967, the International System of Units (SI) has defined the second as the duration of 9192631770 cycles of radiation corresponding to the transition between the two hyperfine ground-state energy levels of the caesium-133 atom. The current clock is accurate to one part in 10^{-16}. For details, see http://www.nist.gov/pml/div688/grp50/primary-frequency-standards.cfm

which means (reading from left to right), the first experiment selects a', but emits in a'', the second experiment selects a'' but emits in a'''. Since everything emitted by the first measurement is accepted by the second measurement, the net effect is to select a' and to emit in a'''. This equation generalizes

$$|a'||a'| = |a'|. \tag{5.40}$$

On the other hand,

$$|a'a''||a''', a^{iv}| = 0 \quad \text{if} \quad a'' \neq a''', \tag{5.41}$$

because what is produced by the first stage cannot enter the second stage; the second rejects what the first emits. This generalizes

$$|a'||a''| = 0, \quad a' \neq a''. \tag{5.42}$$

We can put these two statements together by defining a δ-symbol:

$$\delta(a', a'') = \begin{cases} a' = a'' : 1, \\ a' \neq a'' : 0. \end{cases} \tag{5.43}$$

Then, Eqs. (5.40) and (5.42) are subsumed in

$$|a'||a''| = \delta(a', a'')|a'|, \tag{5.44}$$

and Eqs. (5.39) and (5.41) are combined in

$$|a'a''||a'''a^{iv}| = \delta(a'', a''')|a'a^{iv}|. \tag{5.45}$$

Here's a consistency check: Start from the special case

$$|a'a'||a''a'''| = \delta(a', a'')|a'a'''|, \tag{5.46}$$

and use the construction of the unit symbol, the completely unselective measurement,

$$1 = \sum_{a'} |a'| = \sum_{a'} |a'a'|, \tag{5.47}$$

to see that

$$1|a''a'''| = \sum_{a'} |a'a'||a''a'''| = \sum_{a'} \delta(a', a'')|a'a'''| = |a''a'''|. \tag{5.48}$$

Now, notice we have entered something new:

$$\begin{aligned} |a'a''||a'''a^{iv}| &= \delta(a'', a''')|a'a^{iv}|, &\text{(5.49a)}\\ |a'''a^{iv}||a'a''| &= \delta(a^{iv}, a')|a'''a''|, &\text{(5.49b)} \end{aligned}$$

are not the same:

$$|a'a''||a'''a^{iv}| \neq |a'''a^{iv}||a'a''|. \tag{5.50}$$

Multiplication is not *commutative*. This is obvious from the physical meaning of measurement symbols: If $a' \neq a''$,

$$|a'||a'a''| = |a'a''| \tag{5.51}$$

signifies a measurement in which a' is accepted, but a'' is emitted, while

$$|a'a''||a'| = 0 \tag{5.52}$$

means that everything emited by the first measurement is rejected by the second. Nothing comes out. The order of physical operations is important. Mathematically, this means that the ordering of multiplication of measurement symbols is significant.

5.2 More about spin-1/2

Remember that we defined for spin 1/2

$$J_z = \frac{1}{2}\hbar\sigma_z, \tag{5.53}$$

where the possible values of σ_z, $\sigma_z' = \pm 1$. For this system, there are four possible measurement symbols

$$|\sigma_z'\sigma_z''|, \quad \text{where} \quad \sigma_z' = \pm 1, \sigma_z'' = \pm 1, \tag{5.54}$$

where the two values may be independently assumed. Recall in general that

$$A = \sum_{a'} |a'|, \tag{5.55}$$

so

$$\sigma_z = |+| - |-| = |++| - |--|. \tag{5.56}$$

Note that $\sigma_z^2 = 1$; since $\sigma_z' = \pm 1$, the only possible value of σ_z^2 is 1. (This also follows from $\prod_{a'}(A - a') = 0$.) Check this property explicitly:

$$\sigma_z^2 = (|+| - |-|)(|+| - |-|) = |+||+| - |-||+| - |+||-| + |-||-| = |+| + |-| = 1. \tag{5.57}$$

There is nothing special about the z direction; we must be able to write

$$\mathbf{J} = \frac{1}{2}\hbar\boldsymbol{\sigma}. \tag{5.58}$$

What are σ_x, σ_y? They must have the properties

$$\sigma_x^2 = 1, \quad \sigma_y^2 = 1, \tag{5.59}$$

since nothing can be special about the z axis. Now we know from the general properties of the measurement symbols that

$$|-+||+-| = |--| = |-|, \tag{5.60a}$$
$$|+-||-+| = |++| = |+|, \tag{5.60b}$$

while

$$|-+||-+| = 0 = |+-||+-|. \tag{5.60c}$$

Thus we have

$$(|-+|+|+-|)^2 = |-+||+-|+|+-||-+| = |-|+|+| = 1, \tag{5.61}$$

which is just the property σ_x is supposed to have, so we take

$$\sigma_x = |-+|+|+-|. \tag{5.62}$$

We need an independent combination of measurement symbols, with the same property, to give σ_y. Note that

$$(|-+|-|+-|)^2 = -|-|-|+| = -1, \tag{5.63}$$

since the same cross terms contribute, but with a minus sign, which means that

$$[i(|-+|-|+-|)]^2 = 1, \tag{5.64}$$

so we can adopt

$$\sigma_y = i|-+|-i|+-|. \tag{5.65}$$

To summarize the four combinations of measurement symbols we have found:

$$\sigma_x = |-+|+|+-|, \tag{5.66a}$$
$$\sigma_y = i|-+|-i|+-|, \tag{5.66b}$$
$$\sigma_z = |+|-|-|, \tag{5.66c}$$
$$1 = |+|+|-|. \tag{5.66d}$$

The four symbols $\sigma_{x,y,z}$, 1 are here expressed as linear combinations of measurement symbols. There is something remarkable about how this fits together: There are three dimensions of space, hence 3 σ's, and there are just the right number of measurement symbols to express these.

Now, let's work out the algebra of these σ's:

$$\sigma_x\sigma_y = (|-+|+|+-|)(i|-+|-i|+-|) = i|+|-i|-| = i\sigma_z, \tag{5.67a}$$
$$\sigma_y\sigma_x = (i|-+|-i|+-|)(|-+|+|+-|) = -i|+|+i|-| = -i\sigma_z, \tag{5.67b}$$

from which follows

$$\sigma_x\sigma_y = -\sigma_y\sigma_x, \quad \text{or} \quad \sigma_x\sigma_y + \sigma_y\sigma_x = 0. \tag{5.68}$$

σ_x and σ_y do not commute; we say they *anticommute*. The remaining products are

$$\sigma_y\sigma_z = (i|-+|-i|+-|)(|++|-|--|) = i|-+|+i|+-| = i\sigma_x, \tag{5.69a}$$
$$\sigma_z\sigma_y = (|++|-|--|)(i|-+|-i|+-|) = -i|+-|-i|-+| = -i\sigma_x, \tag{5.69b}$$
$$\sigma_z\sigma_x = (|++|-|--|)(|-+|+|+-|) = |+-|-|-+| = i\sigma_y, \tag{5.69c}$$
$$\sigma_x\sigma_z = (|-+|+|+-|)(|++|-|--|) = |-+|-|+-| = -i\sigma_y. \tag{5.69d}$$

In summary,

$$\begin{aligned} \sigma_x\sigma_y &= i\sigma_z, \quad \sigma_y\sigma_z = i\sigma_x, \quad \sigma_z\sigma_x = i\sigma_y, \\ \sigma_y\sigma_x &= -i\sigma_z, \quad \sigma_z\sigma_y = -i\sigma_x, \quad \sigma_x\sigma_z = -i\sigma_y. \end{aligned} \tag{5.70}$$

Note the cyclic pattern here. A compact notation can be invented for these. If we replace the labels x, y, z by 1, 2, 3, we can write

$$i \neq j: \quad \sigma_i\sigma_j = i\sum_{k=1}^{3}\epsilon_{ijk}\sigma_k \equiv i\epsilon_{ijk}\sigma_k, \tag{5.71}$$

where ϵ_{ijk} is the totally antisymmetric symbol:

$$\epsilon_{123} = +1, \quad \epsilon_{ijk} = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji} = \epsilon_{jki} = \epsilon_{kij}, \tag{5.72}$$

and we have adopted the (Einstein) summation convention for repeated indices. Incorporating also the fact that

$$\sigma_i^2 = 1, \quad i = 1, 2, 3, \tag{5.73}$$

we have

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k, \tag{5.74}$$

where the Kronecker δ symbol is

$$\delta_{ij} = \begin{cases} 1, \ i = j, \\ 0, \ i \neq j. \end{cases} \tag{5.75}$$

Suppose we consider any two vectors

$$\mathbf{a} = (a_1, a_2, a_3), \quad \mathbf{b} = (b_1, b_2, b_3). \tag{5.76}$$

Then Eq. (5.74) implies

$$a_ib_j\sigma_i\sigma_j = a_ib_j\delta_{ij} + i\epsilon_{ijk}a_ib_j\sigma_k, \tag{5.77}$$

or

$$(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{b}) = \mathbf{a}\cdot\mathbf{b} + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b}), \tag{5.78}$$

because

$$\epsilon_{ijk}a_ib_j\sigma_k = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ \sigma_1 & \sigma_2 & \sigma_3 \end{vmatrix} = \mathbf{a}\cdot(\mathbf{b}\times\boldsymbol{\sigma}) = \boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b}). \tag{5.79}$$

5.3 Unitary transformation

If $\mathbf{a} = \mathbf{b} = \mathbf{e}$, where $\mathbf{e}$ is a unit vector, $\mathbf{e}\cdot\mathbf{e} = 1$, Eq. (5.78) means

$$(\boldsymbol{\sigma}\cdot\mathbf{e})^2 = 1. \tag{5.80}$$

If we have three mutually orthogonal unit vectors, forming a right-handed set,

$$\mathbf{e}_i^2 = 1, \quad \mathbf{e}_i \times \mathbf{e}_j = \epsilon_{ijk}\mathbf{e}_k, \tag{5.81}$$

so for example $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, we have, for $i \neq j$,

$$(\boldsymbol{\sigma} \cdot \mathbf{e}_i)(\boldsymbol{\sigma} \cdot \mathbf{e}_j) = i\boldsymbol{\sigma} \cdot (\mathbf{e}_i \times \mathbf{e}_j) = i\epsilon_{ijk}(\boldsymbol{\sigma} \cdot \mathbf{e}_k) \tag{5.82}$$

But these three unit vectors define a coordinate system; say the primed coordinate system:

$$\boldsymbol{\sigma} \cdot \mathbf{e}_i = \sigma_{i'} \quad (\sigma_{x'}, \sigma_{y'}, \sigma_{z'}), \tag{5.83}$$

so in any coordinate system

$$\sigma_{i'}^2, \quad \sigma_{i'}\sigma_{j'} = i\epsilon_{ijk}\sigma_{k'}, \quad i \neq j \tag{5.84}$$

The algebraic properties are independent of the coordinate system.

The two coordinate systems (x, y, z), (x', y', z') differ by a rotation. We want to see in more detail how the new σ's, $\sigma_{x'}$, $\sigma_{y'}$, $\sigma_{z'}$, are related to the original σ's, σ_x, σ_y, σ_z. To do this, we first prove a lemma,

$$\cos\phi + i\sigma_z \sin\phi = e^{i\sigma_z\phi}. \tag{5.85}$$

The proof ot this depends on the definition of a function of an algebraic element,

$$f(A) = \sum_{a'} f(a')|a'|. \tag{5.86}$$

Here

$$|+| = \frac{1+\sigma_z}{2}, \quad |-| = \frac{1-\sigma_z}{2}, \tag{5.87}$$

so

$$\begin{aligned} e^{i\sigma_z\phi} &= e^{i\phi}\frac{1+\sigma_z}{2} + e^{-i\phi}\frac{1-\sigma_z}{2} \\ &= \cos\phi + i\sin\phi\,\sigma_z. \end{aligned} \tag{5.88}$$

Let's also check that this exponential has the expected mulitiplication property.

$$e^{i\phi\sigma_z}e^{i\phi'\sigma_z} = e^{i(\phi+\phi')\sigma_z}. \tag{5.89}$$

Indeed, the left side of this equation is

$$\begin{aligned} \left(e^{i\phi}|+|+e^{-i\phi}|-|\right)\left(e^{i\phi'}|+|+e^{-i\phi'}|-|\right) &= e^{i(\phi+\phi')}|+| = e^{-i(\phi+\phi')}|-| \\ &= e^{i(\phi+\phi')\sigma_z}. \end{aligned} \tag{5.90}$$

This implies

$$e^{-i\phi\sigma_z} = \left(e^{-i\phi\sigma_z}\right)^{-1} \tag{5.91}$$

since

$$e^{-i\phi\sigma_z}e^{i\phi\sigma_z} = 1. \tag{5.92}$$

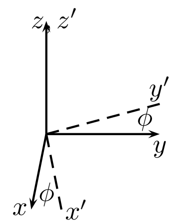

Figure 5.2: Two coordinate systems, (x, y, z), (x', y', z'), which are related by a rotation about the z axis through an angle ϕ.

Let's consider a rotation about the z axis, as shown in Fig. 5.2. So

$$\sigma_{x'} = \sigma_x \cos\phi + \sigma_y \sin\phi, \tag{5.93a}$$
$$\sigma_{y'} = -\sigma_x \sin\phi + \sigma_y \cos\phi, \tag{5.93b}$$
$$\sigma_{z'} = \sigma_z. \tag{5.93c}$$

We know already that

$$\sigma_{x'}^2 = \sigma_{y'}^2 = \sigma_{z'}^2 = 1, \quad \sigma_{x'}\sigma_{y'} = i\sigma_{z'}, \quad \text{etc.} \tag{5.94}$$

But how does this come about? Note that

$$\begin{aligned}\sigma_{x'} &= \sigma_x \cos\phi - i\sigma_z\sigma_x \sin\phi \\ &= (\cos\phi - i\sigma_z \sin\phi)\sigma_x = e^{-i\phi\sigma_z}\sigma_x,\end{aligned} \tag{5.95}$$

or

$$\begin{aligned}\sigma_{x'} &= \sigma_x \cos\phi + i\sigma_x\sigma_z \sin\phi \\ &= \sigma_x(\cos\phi + i\sigma_z \sin\phi) = \sigma_x e^{i\phi\sigma_z}.\end{aligned} \tag{5.96}$$

When we commute σ_x past a function of σ_z, the sign of σ_z changes, since σ_x and σ_z anticommute:

$$\sigma_x\sigma_z = -\sigma_z\sigma_x. \tag{5.97}$$

(Any function of σ_z is actually a linear function:

$$f(\sigma_z) = f(+)\frac{1+\sigma_z}{2} + f(-)\frac{1-\sigma_z}{2} = a + b\sigma_z.) \tag{5.98}$$

If we decompose the exponentials,

$$e^{-i\phi\sigma_z} = e^{-i\frac{\phi}{2}\sigma_z} e^{-i\frac{\phi}{2}\sigma_z}, \tag{5.99a}$$
$$e^{i\phi\sigma_z} = e^{i\frac{\phi}{2}\sigma_z} e^{i\frac{\phi}{2}\sigma_z}, \tag{5.99b}$$

we can write

$$\sigma_{x'} = e^{-i\phi\sigma_z}\sigma_x = e^{-i\frac{\phi}{2}\sigma_z}\sigma_x e^{i\frac{\phi}{2}\sigma_z}, \tag{5.100}$$
$$= \sigma_x e^{i\phi\sigma_z} = e^{-i\frac{\phi}{2}\sigma_z}\sigma_x e^{i\frac{\phi}{2}\sigma_z}, \tag{5.101}$$

in each case commuting σ_x with one of the exponentials. Define

$$U = e^{i\frac{\phi}{2}\sigma_z}; \tag{5.102}$$

it is an algebraic element corresponding to a rotation of the coordinate system about the z axis through an angle ϕ. We have

$$\sigma_{x'} = U^{-1}\sigma_x U. \tag{5.103}$$

It is also true that

$$\sigma_{y'} = U^{-1}\sigma_y U, \tag{5.104a}$$
$$\sigma_{z'} = U^{-1}\sigma_z U. \tag{5.104b}$$

Proof: the last is easy, since σ_z commmutes with itself:

$$\begin{aligned} U^{-1}\sigma_z U &= e^{-i\frac{\phi}{2}\sigma_z}\sigma_z e^{i\frac{\phi}{2}\sigma_z} \\ &= \sigma_z e^{-i\frac{\phi}{2}\sigma_z} e^{i\frac{\phi}{2}\sigma_z} = \sigma_z. \end{aligned} \tag{5.105}$$

Nontrivial is

$$\begin{aligned} U^{-1}\sigma_y U &= e^{-i\frac{\phi}{2}\sigma_z}\sigma_y e^{i\frac{\phi}{2}\sigma_z} = \sigma_y e^{i\phi\sigma_z} \\ &= \sigma_y(\cos\phi + i\sigma_z\sin\phi) = \sigma_y\cos\phi - \sigma_x\sin\phi = \sigma_{y'}, \end{aligned} \tag{5.106}$$

since σ_y, σ_z anticommute, and $\sigma_y\sigma_z = i\sigma_x$.

The discussion here has been restricted to spin 1/2, but we recognize here

$$J_z = \frac{\hbar}{2}\sigma_z, \tag{5.107}$$

so it is plausible (and true) that the symbol which describes (or produces) rotations about the z axis through an angle ϕ for an arbitrary system is

$$U = e^{i\phi J_z/\hbar}. \tag{5.108}$$

Transformations of this type are guaranteed to preserve algebraic relations. Thus, suppose we have elements of our algebra

$$X, Y, Z, W, \quad \text{where} \quad XY = Z, \quad X + Y = W. \tag{5.109}$$

Then if we define transformed elements by

$$\bar{X} = U^{-1}XU, \quad \bar{Y} = U^{-1}YU, \quad \bar{Z} = U^{-1}ZU, \quad \bar{W} = U^{-1}WU, \tag{5.110}$$

then

$$\bar{X} + \bar{Y} = U^{-1}XU + U^{-1}YU = U^{-1}(X+Y)U = U^{-1}WU = \bar{W}, \tag{5.111a}$$

and

$$\bar{X}\bar{Y} = U^{-1}XUU^{-1}YU = U^{-1}XYU = U^{-1}ZU = \bar{Z}, \tag{5.111b}$$

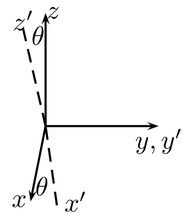

Figure 5.3: Rotation of the Cartesian coordinate system about the y axis through an angle θ.

that is,

$$\bar{X}\bar{Y} = \bar{Z}, \quad \bar{X} + \bar{Y} = \bar{W}, \tag{5.112}$$

which is just the transform of Eq. (5.109). This property is completely independent of what U is.

Therefore, because

$$\sigma_{x'} = U^{-1}\sigma_x U, \quad \sigma_{y'} = U^{-1}\sigma_y U, \quad \sigma_{z'} = U^{-1}\sigma_z U, \tag{5.113}$$

we are guaranteed that $\sigma_{x'}$, $\sigma_{y'}$, $\sigma_{z'}$ satisfy the same algebra as σ_x, σ_y, σ_z. Such transformations are called "unitary transformations," and preserve algebraic structure. Writing the transformation in this way separates the object being rotated ($\boldsymbol{\sigma}$) from the rotation of the coordinate system (represented by U). We must be able to do this in general.

Let's now do a rotation about the y axis, as shown in Fig. 5.3. Geometrically,

$$\sigma_{x'} = \sigma_x \cos\theta - \sigma_z \sin\theta, \tag{5.114a}$$
$$\sigma_{y'} = \sigma_y, \tag{5.114b}$$
$$\sigma_{z'} = \sigma_x \sin\theta + \sigma_z \cos\theta. \tag{5.114c}$$

We expect by analogy that the unitary transformation which does this is given by

$$U = e^{i\frac{\theta}{2}\sigma_y}, \tag{5.115}$$

which represents a rotation about the y axis through the angle θ. This must be so, since all directions are on the same footing. Does it work?

$$\begin{aligned}\sigma_{x'} &= U^{-1}\sigma_x U = e^{-i\frac{\theta}{2}\sigma_y}\sigma_x e^{i\frac{\theta}{2}\sigma_y} = \sigma_x e^{i\theta\sigma_y} \\ &= \sigma_x(\cos\theta + i\sigma_y\sin\theta) = \sigma_x\cos\theta - \sigma_z\sin\theta,\end{aligned} \tag{5.116a}$$

$$\sigma_{y'} = U^{-1}\sigma_y U = e^{-i\frac{\theta}{2}\sigma_y}\sigma_y e^{i\frac{\theta}{2}\sigma_y} = \sigma_y, \tag{5.116b}$$

$$\begin{aligned}\sigma_{z'} &= U^{-1}\sigma_x U = e^{-i\frac{\theta}{2}\sigma_y}\sigma_z e^{i\frac{\theta}{2}\sigma_y} = \sigma_z e^{i\theta\sigma_y} \\ &= \sigma_z(\cos\theta + i\sigma_y\sin\theta) = \sigma_z\cos\theta + \sigma_x\sin\theta,\end{aligned} \tag{5.116c}$$

since σ_y anticommutes with σ_x and σ_z, and

$$e^{i\theta\sigma_y} = \cos\theta + i\sigma_y \sin\theta, \tag{5.117}$$

which is proved just as the corresponding statement for σ_z. All that is essential to the proof is $(i\sigma_y)^2 = -1$.

Quite generally, a rotation about any axis can be constructed in this way. Any rotation is equivalent to a rotation about a fixed axis. Consider a general rotation, so that the z' axis is determined by polar and azimuthal angles, θ and ϕ:

$$\sigma_{z'} = \sigma_x \sin\theta\cos\phi + \sigma_y \sin\theta\sin\phi + \sigma_z\cos\theta, \tag{5.118}$$

as shown in Fig. 4.5. How can this be obtained by a unitary transformation? Note that

$$\begin{aligned}\sigma_{z'} &= \sin\theta e^{-i\frac{\phi}{2}\sigma_z}\sigma_x e^{i\frac{\phi}{2}\sigma_z} + \sigma_z\cos\theta \\ &= e^{-i\frac{\phi}{2}\sigma_z}(\sigma_x\sin\theta + \sigma_z\cos\theta)e^{i\frac{\phi}{2}\sigma_z} \\ &= e^{-i\frac{\phi}{2}\sigma_z}e^{-i\frac{\theta}{2}\sigma_y}\sigma_z e^{i\frac{\theta}{2}\sigma_y}e^{i\frac{\phi}{2}\sigma_z};\end{aligned} \tag{5.119}$$

in effect we undid the ϕ, θ rotations to bring $\sigma_{z'}$ back to the z axis. Let

$$U = e^{i\frac{\theta}{2}\sigma_y}e^{i\frac{\phi}{2}\sigma_z}. \tag{5.120}$$

Then

$$U^{-1} = e^{-i\frac{\phi}{2}\sigma_z}e^{-i\frac{\theta}{2}\sigma_y} \tag{5.121}$$

Note in U^{-1} the factors appear in reverse order. This is a general property with algebraic elements: If X, Y have inverses X^{-1}, Y^{-1}, so

$$X^{-1}X = 1, \quad Y^{-1}Y = 1, \tag{5.122}$$

then the inverse of XY is $(XY)^{-1} = Y^{-1}X^{-1}$ because

$$(Y^{-1}X^{-1})(XY) = Y^{-1}Y = 1. \tag{5.123}$$

Thus we have proved

$$\sigma_{z'} = U^{-1}\sigma_z U, \tag{5.124}$$

for a general rotation of the coordinate system. The rotation is effected in stages: first a rotation about the y axis through an angle θ, then a rotation about the z axis through an angle ϕ. We will see this more completely later on. (See homework.)

To summarize: Independence of directions, or of the orientation of the coordinate system, follows from the non-commutativity of the perpendicular components of the angular momentum. The impossibility of measuring the different components of the angular momentum is now faithfully represented. A small change of the coordinate system does *not* imply a small change in the measured values, which are for spin-1/2 always $\sigma'_{z'} = \pm 1$.