Chapter 2

Axioms and Tools (Part B)

2.1 Axiom 2: Measurement

We now come to one of the most fundamental, yet counterintuitive (because it wildly departs from the classical world) axioms of quantum mechanics, related to how a measurement is performed.

2.1.1 Measurement Outcomes

Quite generally, a measurement is a process in which information about the state of a physical system is acquired by an observer. An observable is a property of a physical system that in principle can be measured. Such property could be for example momentum and spin components, etc. In quantum mechanics, it is postulated that an observable is represented by an Hermitian (also known as self-adjoint) operator acting on the vector space of quantum states. The fundamental axiom of quantum mechanics (that cannot be proven) is that the measurement of an observable $\hat{A}$ prepares an eigenstate of the hermitian operator $\hat{A}$, and the observer learns the value of the corresponding eigenvalue.

In essence, let us assume that we have a certain quantum system described by a ket $|\psi\rangle$, whose expansion in the eigen-kets of $\hat{A}$ reads

$$|\psi\rangle = \sum_i c_i |A_i\rangle \ , \tag{2.1.1}$$

where, as shown before, the expansion coefficients (amplitudes) are $c_i = \langle A_i|\psi\rangle$. We also assume here that the state $|\psi\rangle$ is a ray chosen to be normalized, thus $\sum_i |c_i|^2 = 1$.

The measurement axiom means that when we measure the operator $\hat{A}$, the state $|\psi\rangle$ immediately collapses into one of the possible eigenstates $|A_i\rangle$ of $\hat{A}$, and the result of that specific measurement will be the associated eigenvalue, a_i. The most important aspect of the measurement process is that which of the several eigen-kets is obtained is determined only probabilistically. Specifically, we say that the outcome a_i for the measurement is obtained with a priori probability

$$\text{Prob}(a_i) = |\langle A_i|\psi\rangle|^2 = \langle\psi|A_i\rangle\langle A_i|\psi\rangle = |c_i|^2 \, , \tag{2.1.2}$$

known as Born's probability rule, introduced by Max Born in 1926. We can immediately verify that the probability defined by Born's rule is a correct probability, in the sense that it is correctly normalized:

$$\sum_i \text{Prob}(a_i) = \sum_i |c_i|^2 = \langle\psi|\psi\rangle = 1\,, \tag{2.1.3}$$

where the last equality comes from the normalization condition of the state $|\psi\rangle$.

2.1.2 Repeated measurements

If many identically prepared systems are measured, each described by the same state $|\psi\rangle$, then the expectation value of the outcomes is what you would expect from standard probability theory, namely

$$\begin{aligned}\langle A\rangle \equiv \sum_i a_i \text{Prob}(a_i) &= \\ &= \sum_i a_i \langle\psi|A_i\rangle\langle A_i|\psi\rangle = \\ &= \sum_i \langle\psi|\hat{A}|A_i\rangle\langle A_i|\psi\rangle = \\ &= \langle\psi|\hat{A}\underbrace{\sum_i |A_i\rangle\langle A_i|}_{\hat{1}}|\psi\rangle = \\ &= \langle\psi|\hat{A}|\psi\rangle\ .\end{aligned} \tag{2.1.4}$$

The latter equation is one of the most important equations of quantum mechanics, since it relates the average result for repeated measurements to the quantity $\langle\psi|\hat{A}|\psi\rangle$, known as the "expectation value" of the corresponding operator.

It should be understood here that by "repeated measurement" we mean, strictly, preparing the state $|\psi\rangle$ several times, and each time measuring the observable A. Each observation $k = 1, 2, \ldots, M$ will result in a random result $r_k \in \{a_1, \ldots, a_n\}$. An experimental observer can then estimate $\langle A\rangle$ with the simple mean

$$\langle A\rangle \simeq \frac{1}{M}\sum_k r_k\,, \tag{2.1.5}$$

and in the limit $M \to \infty$ this will coincide with the computed expression $\langle\psi|\hat{A}|\psi\rangle$.

A dramatically different scenario is instead obtained if we prepare the state $|\psi\rangle$ only once, and we perform a measurement over the same state over and over again. In this case, after the first measurement, the state will collapse to a corresponding random eigen-ket, say

$$|\psi\rangle \xrightarrow{\text{measurement}} |A_i\rangle\ , \tag{2.1.6}$$

with probability $P_i = |\langle A_i|\psi\rangle|^2$, and resulting in the value a_i for the measurement outcome. However, since the new state resulting from the measurement is just an eigenstate of the measurement operator, we have that the new state has $c_i' = 1$, thus if we measure the operator $\hat{A}$ again, the result of the measurement will be again a_i, with probability 1, a deterministic measurement!

2.2 Compatible and Incompatible Observables

Two observables $\hat{A}$ and $\hat{B}$ are said to be compatible when their commutator $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$ is vanishing, i.e. $[\hat{A}, \hat{B}] = 0$. Otherwise, if the commutator is non-zero, then these operators are referred to as incompatible. Consider the Pauli matrices σ_x, σ_z defined as:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.2.1}$$

These two matrices (when multiplied by $\hbar/2$) represent physical quantities (the x, z components of the spin operator). It can be shown (exercise) that these two operators $\hat{\sigma}_x, \hat{\sigma}_z$ are incompatible, since

$$[\hat{\sigma}_z, \hat{\sigma}_x] = 2i \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = 2i\hat{\sigma}_y\,. \tag{2.2.2}$$

where $\hat{\sigma}_y$ is the y component of the spin operator. Identifying which operators are compatible and which not has an important consequence, which we will show.

Suppose that $\hat{A}$ and $\hat{B}$ are compatible observables, and the eigenvalues of $\hat{A}$ are non-degenerate. Then the matrix representation of $\hat{B}$ is diagonal in the basis of A, thus $\langle A_i|\hat{B}|A_j\rangle = \delta_{ij} b_i$, where the lower script A here denotes that the matrix elements are with respect to the eigenkets of A. Let us prove this claim. Using the fact that $\hat{A}$ and $\hat{B}$ commute, we have

$$\begin{aligned} \langle A_i|[\hat{A}, \hat{B}]|A_j\rangle &= \langle A_i|\hat{A}\hat{B} - \hat{B}\hat{A}|A_j\rangle = \\ &= (a_i - a_j)\,\langle A_i|\hat{B}|A_j\rangle = \\ &= 0\,, \end{aligned} \tag{2.2.3}$$

thus $\langle A_i|\hat{B}|A_j\rangle$ must vanish for $i \neq j$.

Importantly, from this proof we find that $|A_i\rangle$ and $|B_i\rangle$ are eigen-kets of both $\hat{A}$ and $\hat{B}$. This can shown using the fact that $\hat{B}$ is diagonal in the $\hat{A}$-eigenbasis, thus the decomposition

$$\hat{B} = \sum_i b_i\,|A_i\rangle\,\langle A_i|\;, \tag{2.2.4}$$

holds, which immediately implies that if we apply this operator to an eigen-ket of $\hat{A}$, we get

$$\begin{aligned} \hat{B}\,|A_j\rangle &= \sum_i b_i\,|A_i\rangle\,\langle A_i|A_j\rangle = \\ &= b_j\,|A_j\rangle\;. \end{aligned} \tag{2.2.5}$$

Thus, $|A_j\rangle$ is an eigen-ket of $\hat{B}$, and we also identify the diagonal matrix elements as the eigenvalues b_j. In general, when a certain ket is an eigen-ket of more than one operator, we typically denote it as $|A_i, B_i\rangle$ or, often, with a collective name $|K_i\rangle$.

Fundamental consequence of the commutativity of observables is therefore that the measurement process in this case is familiarly similar to what would happen in the classical case. For example, imagine again a state $|\psi\rangle$, and that we measure the observable A, then the result of the measurement will yield some random value a_j and the state will collapse into the corresponding eigenstate $|A_j\rangle$. Measuring now B will result in the value b_j, with probability 1 (recall that $|A_j\rangle$ is also an eigenvalue of B). Further measuring A would again return a_j, thus the measurement done with B has not destroyed (or affected in any way) the state of the system, as per the observable B is concerned. This is a familiar situation in classical mechanics, in the sense that we can expect to be able to measure different quantities (say, velocity and position of a particle, for example) without changing the state of the system itself. This notion however breaks dramatically when considering non-commuting observables.

Non-commuting observables $\hat{A}, \hat{B}$ satisfy the commutation relation $[\hat{A}, \hat{B}] \neq 0$. In this case, $\hat{A}$ and $\hat{B}$ do not share a set of common eigen-kets in general. Let us proof this claim. Suppose that the converse is true, i.e. that we can find a set of common eigen-kets. Then we have that

$$\hat{A}\hat{B}\,|A_i, B_i\rangle = \hat{A}b_i\,|A_i, B_i\rangle = a_i b_i\,|A_i, B_i\rangle \tag{2.2.6}$$

and also

$$\hat{B}\hat{A}\,|A_i, B_i\rangle = \hat{B}a_i\,|A_i, B_i\rangle = a_i b_i\,|A_i, B_i\rangle \ . \tag{2.2.7}$$

Since $\hat{A}\hat{B}\,|A_i, B_i\rangle = \hat{B}\hat{A}\,|A_i, B_i\rangle$ for all eigen-kets i, then this implies $\hat{A}\hat{B} = \hat{B}\hat{A}$, which is in contradiction of the assumption.

2.3 The Uncertainty Principle

In Subsection 2.1.2, we have analyzed the case of repeated measurements, and came to the conclusion that expectation value of a given operator over many experiments is given by:

$$\langle A\rangle \equiv \langle\psi|\hat{A}|\psi\rangle \ . \tag{2.3.1}$$

In addition to the expectation value, we can also compute the variance associated with the measurement of the operator. To this end, we introduce the displacement operator defined as:

$$\Delta\hat{A} = \hat{A} - \langle A\rangle\,\hat{I} \,, \tag{2.3.2}$$

such that the expectation value of its square is the variance:

$$\left\langle \Delta A^2 \right\rangle \equiv \langle\psi|\left(\hat{A} - \langle A\rangle\,\hat{I}\right)^2|\psi\rangle = \langle\psi|\hat{A}^2|\psi\rangle - \langle\psi|\hat{A}|\psi\rangle^2 \ . \tag{2.3.3}$$

This variance really measures how much the outcome of a given measurement is different from its average, exactly following the definition of variance in statistics. For example, if you take the case in which $|\psi\rangle$ is an eigenstate of $\hat{A}$ (say $|A_i\rangle$), it easy to see that

$$\left\langle \Delta A^2 \right\rangle \equiv \langle A_i|\hat{A}^2|A_i\rangle - \langle A_i|\hat{A}|A_i\rangle^2 = a_i^2 - (a_i)^2 = 0 \,, \tag{2.3.4}$$

thus we recover the fundamental measurement postulate, telling us that if we repeatedly measure an eigenstate, we always find the same result (we have zero variance).

The uncertainty principle is an important result connecting the amount of intrinsic uncertainty (variance) associated with the measurement of two observables. It states that, for two observables $\hat{A}$ and $\hat{B}$, we have:

$$\left\langle \Delta A^2 \right\rangle \left\langle \Delta B^2 \right\rangle \geq \frac{1}{4} \left| \left\langle [\hat{A}, \hat{B}] \right\rangle \right|^2 . \tag{2.3.5}$$

Before proving it, let us discuss the consequences of this inequality. There are two cases:

1. The two observables commute, thus $[\hat{A}, \hat{B}] = 0$. In this case, then there is no intrinsic limit on the *precision* we can attain when measuring the two observables on the same state. $\langle \Delta A^2 | \Delta A^2 \rangle$ and $\langle \Delta B^2 | \Delta B^2 \rangle$ can be as small as we want.

2. The two observables do not commute, thus $[\hat{A}, \hat{B}] \neq 0$. In this case, there is an intrinsic limit on the precision we can attain when measuring the two observables.

In order to prove Eq. (2.3.5), we first need two intermediate results. Firstly, the Cauchy–Schwarz inequality:

$$\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle \geq | \langle \alpha | \beta \rangle |^2, \tag{2.3.6}$$

which is a generalization of the triangle inequality to other metric spaces with a given inner product. Let us proof this inequality.

The inequality can be proven in a variety of ways. Here we consider the ket

$$|C_\lambda\rangle = |\alpha\rangle + \lambda |\beta\rangle , \tag{2.3.7}$$

obtained as a linear combination, with complex λ, of the two given kets. The norm of this ket is obviously positive, thus

$$\begin{aligned} \langle C_\lambda | C_\lambda \rangle &= \langle \alpha | \alpha \rangle + \langle \beta | \beta \rangle |\lambda|^2 + \langle \alpha | \beta \rangle \lambda^* + \langle \beta | \alpha \rangle \lambda \\ &\geq 0 . \end{aligned}$$

This inequality holds for all values of λ, and the Cauchy–Schwarz inequality is found considering $\lambda = - \langle \alpha | \beta \rangle / \langle \beta | \beta \rangle$, since we have

$$\langle \alpha | \alpha \rangle - \frac{| \langle \alpha | \beta \rangle |^2}{\langle \beta | \beta \rangle} \geq 0, \tag{2.3.8}$$

which proves the original inequality (notice that the case in which $\langle \beta | \beta \rangle = 0$ can be easily proven separately).
Using the Cauchy–Schwarz inequality, with

$$|\alpha\rangle = \Delta\hat{A} |\psi\rangle , \quad |\beta\rangle = \Delta\hat{B} |\psi\rangle \tag{2.3.9}$$

we get

$$\left\langle \Delta A^2 \middle| \Delta A^2 \right\rangle \left\langle \Delta B^2 \middle| \Delta B^2 \right\rangle \geq | \langle \psi | \Delta\hat{A} \Delta\hat{B} | \psi \rangle |^2. \tag{2.3.10}$$

where we have used the fact that the displacement operators are Hermitian. We now evaluate the R.H.S. noticing that for two arbitrary Hermitian operators we have

$$\hat{O}_1\hat{O}_2 = \frac{1}{2}[\hat{O}_1, \hat{O}_2] + \frac{1}{2}\{\hat{O}_1, \hat{O}_2\}. \tag{2.3.11}$$

We also notice that

$$\begin{aligned}
\mathrm{Re}\left(\left\langle\left[\hat{O}_1, \hat{O}_2\right]\right\rangle\right) &= \frac{1}{2}\left(\left\langle\left[\hat{O}_1, \hat{O}_2\right]\right\rangle + \left\langle\left[\hat{O}_1, \hat{O}_2\right]\right\rangle^*\right) = \\
&= \frac{1}{2}\left(\left\langle\hat{O}_1\middle|\hat{O}_2\right\rangle - \left\langle\hat{O}_2\middle|\hat{O}_1\right\rangle + \left\langle\hat{O}_1\middle|\hat{O}_2\right\rangle^* - \left\langle\hat{O}_2\middle|\hat{O}_1\right\rangle^*\right) = \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Im}\left(\left\langle\left\{\hat{O}_1, \hat{O}_2\right\}\right\rangle\right) &= \frac{1}{2i}\left(\left\langle\left\{\hat{O}_1, \hat{O}_2\right\}\right\rangle - \left\langle\left\{\hat{O}_1, \hat{O}_2\right\}\right\rangle^*\right) = \\
&= \frac{1}{2i}\left(\left\langle\hat{O}_1\middle|\hat{O}_2\right\rangle + \left\langle\hat{O}_2\middle|\hat{O}_1\right\rangle - \left\langle\hat{O}_1\middle|\hat{O}_2\right\rangle^* - \left\langle\hat{O}_2\middle|\hat{O}_1\right\rangle^*\right) = \\
&= 0,
\end{aligned}$$

thus

$$\begin{aligned}
|\left\langle\psi\middle|\Delta\hat{B}\Delta\hat{A}\middle|\psi\right\rangle|^2 &= \frac{1}{4}\left\{\left|\left\langle[\Delta\hat{A}, \Delta\hat{B}]\right\rangle\right|^2 + \left|\left\langle\{\Delta\hat{A}, \Delta\hat{B}\}\right\rangle\right|^2\right\} = \\
&= \frac{1}{4}\left\{\left|\left\langle[\hat{A}, \hat{B}]\right\rangle\right|^2 + \left|\left\langle\{\Delta\hat{A}, \Delta\hat{B}\}\right\rangle\right|^2\right\},
\end{aligned} \tag{2.3.12}$$

thus omitting the second term, we get the uncertainty inequality.

2.4 Change of Basis

Non-commuting operators define a set of distinct eigen-kets and eigenvalues that can be independently used to describe the same physical system. For example, consider two operators $\hat{A}$ and $\hat{B}$ with eigenvalues a_i, b_i and eigen-kets $|A_i\rangle, |B_i\rangle$. We therefore have the usual eigenvalue relations:

$$\hat{A}\,|A_i\rangle = a_i\,|A_i\rangle\,, \quad \hat{B}\,|B_i\rangle = b_i\,|B_i\rangle \tag{2.4.1}$$

with the orthonormality conditions

$$\langle A_i|A_j\rangle = \delta_{ij}, \quad \langle B_i|B_j\rangle = \delta_{ij}\,. \tag{2.4.2}$$

Since they both form a complete basis for our vector space, it means that an arbitrary ket state can be written in the either of the two orthonormal bases $\{|A_j\rangle\}$ and $\{|B_j\rangle\}$ as follows

$$|\psi\rangle = \sum_j |A_j\rangle\,\langle A_j|\psi\rangle = \sum_j c_j\,|A_j\rangle \tag{2.4.3}$$

$$|\psi\rangle = \sum_j |B_j\rangle\,\langle B_j|\psi\rangle = \sum_j d_j\,|B_j\rangle\,. \tag{2.4.4}$$

Notice that the physical state here is exactly the same, what changes is just how we are mathematically representing it. In quantum physics, it is very often necessary to relate different representations, thus we require a map from one basis to the other. This is again something that is standard in linear algebra, however it is worth recalling it here using the bra-ket formalism.

The main result is that there is a unitary operator $\hat{U}$ that connects the two representations. More specifically,

$$|B_i\rangle = \hat{U} |A_i\rangle \ , \tag{2.4.5}$$

and unitarity here means that

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{\mathbb{1}} \, . \tag{2.4.6}$$

The operator $\hat{U}$ in bra-ket notation takes a very elegant form

$$\hat{U} = \sum_j |B_j\rangle \langle A_j| \ , \tag{2.4.7}$$

which can be verified computing the explicit action of this operator on both the eigenstates of $\hat{A}$ and $\hat{B}$. For example, we have:

$$\hat{U} |A_i\rangle = \sum_j |B_j\rangle \langle A_j|A_i\rangle = |B_i\rangle \ , \tag{2.4.8}$$

and the inverse transformation is found using the conjugate operator

$$\hat{U}^\dagger |B_i\rangle = \sum_j |A_j\rangle \langle B_j|B_i\rangle = |A_i\rangle \ , \tag{2.4.9}$$

where in both cases we have used the orthonormality conditions.
It is also straightforward to verify that the operator is unitary, using the completeness relations for $\hat{B}$:

$$\begin{aligned} \hat{U}\hat{U}^\dagger &= \left(\sum_j |B_j\rangle \langle A_j|\right) \left(\sum_k |A_k\rangle \langle B_k|\right) \\ &= \sum_{jk} |B_j\rangle \langle A_j|A_k\rangle \langle B_k| \\ &= \sum_j |B_j\rangle \langle B_j| \\ &= \hat{I} \, , \end{aligned} \tag{2.4.10}$$

and $\hat{A}$:

$$\begin{aligned} \hat{U}^\dagger\hat{U} &= \left(\sum_j |A_j\rangle \langle B_j|\right) \left(\sum_k |B_k\rangle \langle A_k|\right) \\ &= \sum_{jk} |A_j\rangle \langle B_j|B_k\rangle \langle A_k| \\ &= \sum_j |A_j\rangle \langle A_j| \\ &= \hat{I} \, . \end{aligned} \tag{2.4.11}$$

In the basis of $\hat{A}$, we can also work out the matrix elements of the operator $\hat{U}$:

$$\langle A_i|\hat{U}|A_j\rangle = \langle A_i|\left(\sum_k |B_k\rangle \langle A_k|\right)|A_j\rangle = \langle A_i|B_j\rangle \,, \tag{2.4.12}$$

there the matrix elements are just the amplitudes, or scalar product, between the two sets of eigenstates. Similarly, the matrix elements of $\hat{U}^\dagger$ in the $\hat{B}$ basis are

$$\langle B_i|\hat{U}^\dagger|B_j\rangle = \langle B_i|\left(\sum_k |A_k\rangle \langle B_k|\right)|B_j\rangle = \langle B_i|A_j\rangle \,, \tag{2.4.13}$$

thus

$$\langle A_i|\hat{U}|A_j\rangle = \left(\langle B_j|\hat{U}^\dagger|B_i\rangle\right)^* \,. \tag{2.4.14}$$

2.4.1 Transforming States

Given an explicit form for the transformation matrix, $\hat{U}$, we are therefore in position to solve the problem of finding the coefficients (amplitudes) of a given state in a certain basis (say, $|B_i\rangle$), once its coefficients in another basis, say $|A_i\rangle$, are known. We have

$$\begin{aligned} d_j &= \langle B_j|\psi\rangle = \\ &= \left\langle B_j \middle| \left(\sum_k c_k |A_k\rangle\right)\right\rangle = \\ &= \sum_k \langle B_j|A_k\rangle c_k = \\ &= \sum_k \langle B_j|\hat{U}^\dagger|B_k\rangle c_k. \end{aligned}$$

2.4.2 Transforming Operators

Similar rules can be derived to transform matrix elements of operators, when passing from one representation to another. For example, we can derive rules to obtain matrix elements of a certain operator in two different bases:

$$\langle A_i|\hat{O}|A_j\rangle \to \langle B_i|\hat{O}|B_j\rangle \,.$$

This is (we insert 2 identities):

$$\begin{aligned} \langle B_i|\hat{O}|B_j\rangle &= \langle B_i|\sum_l |A_l\rangle \langle A_l|\,\hat{O}\sum_k |A_k\rangle \langle A_k|B_j\rangle = \\ &= \sum_{lk} \langle B_i|A_l\rangle \, \langle A_l|\hat{O}|A_k\rangle \, \langle A_k|B_j\rangle = \\ &= \sum_{lk} \langle A_i|\hat{U}^\dagger|A_l\rangle \, \langle A_l|\hat{O}|A_k\rangle \, \langle A_k|\hat{U}|A_j\rangle \,, \end{aligned} \tag{2.4.15}$$

which can be written also as a matrix multiplication

$$\hat{O}_{(B)} = \hat{U}^\dagger_{(A)}\hat{O}_{(A)}\hat{U}_{(A)}. \tag{2.4.16}$$

2.4.3 What's changing? States or operators?

We have seen that the unitary matrix corresponding to the rotation operator $\hat{U}$ can be used either to transform *states* or to transform *operators*. The two points of view go as follows:

1. On one side, there is only one observable ($\hat{A}$) and measurements of some other observable (say, $\hat{B}$) are obtained *rotating* the state vector to the corresponding new references basis. In this case then one first prepares a state $|\psi_B\rangle = \hat{U}|\psi\rangle$, and then measures $\hat{A}$ on the new state, such that

$$\langle B\rangle = \langle\psi_B|\hat{A}|\psi_B\rangle\,. \tag{2.4.17}$$

2. On the other side, there is only one physical state ($|\psi\rangle$) and measurements of different observables, for example of $\hat{B}$, are found applying distinct measurement operators on the reference state. Thus we have $\hat{B} = \hat{U}^\dagger \hat{A}\hat{U}$, and

$$\langle B\rangle = \langle\psi|\hat{B}|\psi\rangle\,. \tag{2.4.18}$$

The two formulations are completely equivalent, and it is often the application that tells us which way of thinking makes solving a certain problem easier. When studying time evolution, we will see more prominently what the differences between these two formulations can give us in terms of intuition on the physical systems. In that context, we will identify the first viewpoint as Schrödinger's view on quantum physics, and the second one as Heisenberg's viewpoint.

2.5 Analysis of the Stern–Gerlach experiments

Having introduced the most fundamental postulates of quantum mechanics, we are now in position to resolve one unsatisfactory argument that we had to introduce at the beginning of these lectures: the form of the eigenstates of the spin operators. First of all, it is very important to realize that the kind of experiment we are analyzing here falls under the second type described above (what we have called, repeated measurements). The reason is that the SG apparatus does not measure the spin of an individual electron, but rather of a large number of electrons at once. Schematically, we can think that independent electrons pass through the analyzer and each of them is deflected either upwards or downwards. What we are doing then is essentially equivalent to preparing the same state and measuring many times.

2.5.1 The operator $\hat{S}_z$

When we measure the z component of the spin in our SG apparatus, we postulate the existence of a corresponding measure operator that we call $\hat{S}_z$. When the spin is measured along the z direction, the system will then immediately collapse in one of the two eigenkets of this operator: $|+\rangle$ or $|-\rangle$, and the result of the measurement will be, respectively, one of the two eigenvalues of the $\hat{S}_z$ operator, thus either $a_+ = +\hbar/2$ or $a_- = -\hbar/2$. In

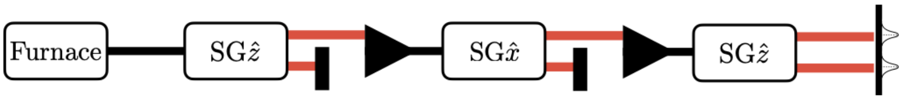

Figure 2.1: Experiment with two orthogonal SG$\hat{z}$ and SG$\hat{x}$ filters each blocking their down streams, followed by a SG$\hat{z}$ device. The last device outputs two streams of equal intensity.

order to find an explicit expression for the spin operators, we start observing that the identity operator for a 2-dimensional finite vector space can be written as

$$\hat{1} = |+\rangle \langle+| + |-\rangle \langle-| \,, \tag{2.5.1}$$

and the representation of $\hat{S}_z$ in the basis of its eigenvectors is just

$$\hat{S}_z = \frac{\hbar}{2} \left(|+\rangle \langle+| - |-\rangle \langle-| \right) . \tag{2.5.2}$$

The operator is also just a 2×2 diagonal matrix in this basis

$$\hat{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{2.5.3}$$

and the notion of eigen-kets being *vectors* is particularly clear when writing them as the algebraic eigenvectors of this matrix, namely

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{2.5.4}$$

This also clarifies why we have previously remarked that kets are column vectors, whereas bras are row vectors:

$$\langle+| = \begin{pmatrix} 1 & 0 \end{pmatrix} , \quad \langle-| = \begin{pmatrix} 0 & 1 \end{pmatrix} . \tag{2.5.5}$$

2.5.2 The operator $\hat{S}_x$

From Figure 2.1, we see that when a beam of type $|S_x;+\rangle$ goes again through a S_z measurement, the beam is deflected in both directions, and we that the counts observed in the SG experiment are equal: $N(+) = N(-)$. This means that, in general,

$$|S_x;+\rangle = \frac{1}{\sqrt{2}} |S_z;+\rangle + \frac{1}{\sqrt{2}} e^{i\delta_1} |S_z;-\rangle \,, \tag{2.5.6}$$

where δ_1 is a real-valued phase that we will determine in a moment. You can verify that this form is correct because:

1. When measuring S_z in the last stage, we apply the operator $\hat{S}_z$, thus the state collapses to one of the two eigenstates of $\hat{S}_z$, with equal probabilities $P(+) = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}$ and $P(-) = \left|\frac{1}{\sqrt{2}} e^{i\delta_1}\right|^2 = \frac{1}{2}$.

Figure 2.2: Experiment with two SG devices aligned in the x and y directions, respectively and blocking one component before entering the second device. All atoms emerge with the unfiltered state. Notice that in this case atoms exiting the furnace arrive from the z direction.

2. The state of Eq. (2.5.6) is a correctly normalized ket, indeed: $\langle S_x;+|S_x;+\rangle = 1$, as it is easy to verify.

We can also find $|S_x;-\rangle$ only using the postulates we have introduced above. Indeed, we know that different eigen-kets of the same operator are orthogonal, thus we must have $\langle S_x;+|S_x;-\rangle = 0$, as well as $\langle S_x;-|S_x;-\rangle = 1$, these two conditions fix the form of the other eigen-ket:

$$|S_x;-\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle - \frac{1}{\sqrt{2}}e^{i\delta_1}|S_z;-\rangle\,. \tag{2.5.7}$$

Now, it should be remarked that using only the results of Experiment 3, we cannot determine the value of the phase factor δ_1, since the information we have from the counts allows us only to reconstruct the square modulus of the amplitudes, and not the amplitudes themselves. To find δ_1 we need more information.

2.5.3 The operator $\hat{S}_y$

A similar analysis can be carried out for the y component of the spin, leading to:

$$|S_y;\pm\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle \pm \frac{1}{\sqrt{2}}e^{i\delta_2}|S_z;-\rangle\,, \tag{2.5.8}$$

where we have introduced yet another phase factor, δ_2, to be determined. In order to determine both phase factors, we consider the experiment as in Figure 2.2. This class of experiments gives the same that was found before when considering the z and x directions, and not surprisingly so, because of symmetry reasons. This however implies that

$$|\langle S_x;+|S_y;+\rangle|^2 = |\langle S_x;+|S_y;-\rangle|^2 = \frac{1}{2}, \tag{2.5.9}$$

when blocking the $x-$ component of the spin (as in Figure) and

$$|\langle S_x;-|S_y;+\rangle|^2 = |\langle S_x;-|S_y;-\rangle|^2 = \frac{1}{2}, \tag{2.5.10}$$

when blocking the $x+$ component of the spin. These conditions allow to fix the phase factors, indeed we find the condition

$$
\begin{aligned}
|\langle S_x;\pm|S_y;+\rangle|^2 &= \left|\left(\frac{1}{\sqrt{2}}\langle S_z;+| \pm \frac{1}{\sqrt{2}}e^{-i\delta_1}\langle S_z;-|\right)\times\left(\frac{1}{\sqrt{2}}|S_z;+\rangle + \frac{1}{\sqrt{2}}e^{i\delta_2}|S_z;-\rangle\right)\right|^2 \\
&= \left|\left(\frac{1}{2}\pm\frac{1}{2}e^{i(\delta_2-\delta_1)}\right)\right|^2 \\
&= \left(\frac{1}{2}\pm\frac{1}{2}e^{i(\delta_2-\delta_1)}\right)\left(\frac{1}{2}\pm\frac{1}{2}e^{-i(\delta_2-\delta_1)}\right) \\
&= \frac{1}{2}. \qquad (2.5.11)
\end{aligned}
$$

This has a solution for $\delta_2-\delta_1=\pm\pi/2$, since $e^{\pm i\frac{\pi}{2}}=\pm i$, thus $\left|\frac{1}{2}(1\pm i)\right|^2=\frac{1}{2}$. While the phase difference is physical, there is no way (but for conventional reasons) to fix separately δ_1 and δ_2. The usual convention is to take $\delta_1=0$ and $\delta_2=\pi/2$, yielding:

$$
|S_x;\pm\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle \pm \frac{1}{\sqrt{2}}|S_z;-\rangle\,, \qquad (2.5.12)
$$

$$
|S_y;\pm\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle \pm i\frac{1}{\sqrt{2}}|S_z;-\rangle\,. \qquad (2.5.13)
$$

The corresponding operators are

$$
\begin{aligned}
\hat{S}_x &= \frac{\hbar}{2}\left(|S_x;+\rangle\langle S_x;+| - |S_x;-\rangle\langle S_x;-|\right) \\
&= \frac{\hbar}{2}\left(|S_z;+\rangle\langle S_z;-| + |S_z;-\rangle\langle S_z;+|\right)\,, \qquad (2.5.14)
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{S}_y &= \frac{\hbar}{2}\left(|S_y;+\rangle\langle S_y;+| - |S_y;-\rangle\langle S_y;-|\right) \\
&= \frac{\hbar}{2}i\left(-|S_z;+\rangle\langle S_z;-| + |S_z;-\rangle\langle S_z;+|\right)\,. \qquad (2.5.15)
\end{aligned}
$$

2.5.4 Pauli matrices

Putting together the results we have found for the three spin operators, we can finally also compute the corresponding matrix representation of these operators. They read:

$$
\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad (2.5.16)
$$

$$
\hat{S}_y = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \qquad (2.5.17)
$$

$$
\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \qquad (2.5.18)
$$

The matrices that appear in these equations are very famous, and called Pauli matrices $\sigma_x,\sigma_y,\sigma_z$, such that $\hat{S}_\alpha=\frac{\hbar}{2}\sigma_\alpha$.

2.5.5 Measurements in different spin bases

We can compute the matrix elements of the transformation matrix connecting the $\hat{S}_z$ basis to the $\hat{S}_x$ basis. The transformation operator is

$$\hat{H} = \sum_{i=\pm} |S_x; i\rangle \langle S_z; i| \,, \tag{2.5.19}$$

and recalling that $|S_x; \pm\rangle = \frac{1}{\sqrt{2}} |S_z; +\rangle \pm \frac{1}{\sqrt{2}} |S_z; -\rangle$, we have that the transformation matrix has the following matrix elements in the $\hat{S}_z$ basis:

$$\hat{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} . \tag{2.5.20}$$

This transformation matrix is of fundamental importance, for example, in quantum computing and it is also known there as *Hadamard gate*. It has the property $\hat{H}\hat{H}^\dagger = \hat{I}$, as we expect from the general properties of the transformation matrices. Measuring the $\hat{S}_x$ operator can be done then in one of two mathematically equivalent ways:

1. There is only one spin operator ($\hat{S}_z$) and measurements of the x component of the spin are found rotating the state vector to the corresponding new references basis[1]. In this case then one prepares a state $|\psi_x\rangle = \hat{H} |\psi\rangle$, and then measures $\hat{S}_z$ on the new state, such that
$$\langle S_x \rangle = \langle \psi_x | \hat{S}_z | \psi_x \rangle \,.$$

2. There is only one state ($|\psi\rangle$) and measurements of different observables, for example of $\hat{S}_x$, are found applying distinct measurement operators on the reference state. Thus we have $\hat{S}_x = \hat{H}^\dagger \hat{S}_z \hat{H}$, as it can be easily checked, and
$$\langle S_x \rangle = \langle \psi | \hat{S}_x | \psi \rangle \,.$$

[1]Notice that rotating to the x basis, which by definition is a rotation in an *abstract vector space*, in this special case corresponds to an actual *physical rotation* in real space of the SG device. This is a quite special coincidence that happens almost exclusively for spin observables.