Definition 13 (Double Slit with measurement)**.** Consider the following variants of the experiment as in Definition 10.

1. Instead of a light source, a single matter particle, such as an electron, is used and its position on the screen detected.
2. Same as 1, except two detectors are placed on the slits (one for each slit).

The idea is to use single particles and try to see which slit the particle goes through. Experiment 2 does exactly this and we use Experiment 1 as a control. So, what does Nature do?

Fact 14. *When Experiment 1 in Definition 13 is performed repeatedly, the particles accumulate such that an interference pattern emerges. When Experiment 2 in Definition 13 is performed, exactly one of the detectors detects a particle (or 'clicks'). However, the interference pattern is replaced with particles accumulating around two peaks, one corresponding to each slit.*

Somehow, the very act of measurement, is disturbing the system. Knowing which slit the particle went through, washes away the interference pattern. Measurements therefore occupy a central role in quantum mechanics.

We end this discussion by returning to the question of what the 'wave' for an electron is. Deferring the formal details, this 'wave' denotes what is known as a 'probability amplitude'. The word amplitude means that this quantity can be both positive and negative (in fact can be a complex number) and thus it can interfere, just as the amplitude of a wave can interfere—a trough and a crest cancel each other. Schrödinger's equation governs how this probability amplitude changes over time. The square of this probability amplitude yields the probability of seeing the particle at any given location.

So an electron beam in double slit experiment may be thought of as a plain wave describing the probability amplitude of the electron, going through the double slits just as a wave does, produces an interference pattern and the probability that the electron is observed at any given location on the screen is then governed by the square of (the absolute value of) this probability amplitude.

§ 2.3 Axioms

2.3.1 Single system

That was a lot of words but all these observations can be explained by just four mathematically precise axioms. The benefit of introducing two concrete experiments was that it will allow us to make sure the abstract axioms make physical sense.

Recall the Silver atom from the Stern Gerlach experiment. A Silver atom can be thought of as a system whose state determines how it behaves as it interacts with other physical objects, such as the Stern Gerlach apparatus. In quantum mechanics, the state of any system, is specified by a vector which is written as $|\psi\rangle$.

Informally stated, quantum mechanics postulate that this state can change in two distinct ways.

1. Unitary evolution. One can apply external forces and internal interactions within parts of the system. These are captured by what is termed a 'Hamiltonion' H of the system. Given the Hamiltonian and the initial state of the system, Schrodinger's equation governs the state at any later time.
2. Measurement. One can choose to observe some property of the system. Suppose an observable O takes values in $\{0,1\}$. A measurement of O when the system is in the state $|\psi\rangle$ results in, in general, the value $b \in \{0,1\}$ with probability p_b and the state of the quantum state after the measurement changes to (or more dramatically, 'collapses to') $|\psi_b\rangle$ where both p_b and $|\psi_b\rangle$ can be computed using b, O and $|\psi\rangle$.

This situation is very unusual, and also somewhat unsatisfactory for the following reasons:

(i) Inherent randomness. In (classical) physics, probabilities only arise when one models the ignorance of the observer. Given a perfect observer who knows the exact initial conditions, (classical) physics is fully deterministic and hence allows the observer to compute exactly what happens at any given later time.[2] Hence, it is very strange that quantum mechanics, at a fundamental level, introduces probabilities—even when the initial state is specified precisely (i.e. the observer knows everything possible about the system).

(ii) The measurement problem. How does Nature know whether to perform a unitary evolution or perform a measurement? Shouldn't a measurement also be governed by some kind of unitary evolution since measurement itself is a physical process involving an observer (and their apparatus) and the system of interest—both of which, by universality

[2] There is something to be said about stability and chaos washing away any predictive power but let us focus on systems that don't have these issues.

of physics, must be described using the same physics. Somehow, there is a divide between the classical world and the quantum world. There are many competing ideas on how to resolve the measurement problem—none of them, to my knowledge, have been experimentally validated (many are 'interpretations of quantum mechanics' in that their predictions are identical to those of quantum mechanics). An interesting reference on the topic is the book titled Something Deeply Hidden by Sean Carrol.

Irrespective of these foundational issues, it turns out that in practice, it is always clear how to apply the postulates of quantum mechanics to obtain quantitative predictions and so far, no deviation from these predictions has been observed in any experiment.

We now turn to the task of formally stating the axioms of quantum mechanics.

We start by setting up some notation and recalling some basic facts about linear algebra. The notation $|\cdot\rangle$ that we used earlier for a vector, is part of the so-called the *Bra-Ket notation*, introduced by Paul Dirac to unify Schrodinger's and Heisenberg's versions of quantum mechanics termed wave mechanics and matrix mechanics respectively. (Interestingly, in the early days, two seemingly different versions of quantum mechanics were discovered). Since physics involves positions and momenta, which are continuous variables, the corresponding vectors become infinite dimensional and this notation is particularly helpful in that context. However, we will restrict ourselves to finite dimensions and even there, the bra-ket notation has now become standard in the field.

Notation 15 (Bra-Ket Notation, Hermitian and Unitary Matrices). For finite dimensional systems, we introduce the following notation.

Bras and Kets (Vectors) Let $\mathcal{H}$ be a complex vector space $\mathbb{C}^d$, of dimension d. Then, a *ket* $|\psi\rangle$ in $\mathcal{H}$ is a sequence of d complex numbers written as a column vector

$$|\psi\rangle = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{bmatrix}$$

and the Hermitian conjugate of $|\psi\rangle$ (or simply the transpose and entry-wise complex conjugate) is denoted by the *bra* $\langle\psi|$ which in turn is written as the row vector

$$\langle\psi| = [\alpha_1^*, \alpha_2^*, \dots \alpha_d^*] .$$

Finally, the inner product between the bra $\langle\psi|$ and a ket

$$|\phi\rangle = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_d \end{bmatrix}$$

is given by $\langle\psi|\phi\rangle := \sum_{i=1}^d \alpha_i^* \beta_i$.

Basis and orthonormal vectors We say that a set of kets $\{|\psi_1\rangle, \dots |\psi_d\rangle\}$ is a *basis* of the vector space $\mathcal{H}$ if every vector $|\psi'\rangle \in \mathcal{H}$ can be uniquely expressed as a linear combination of $\{|\psi_i\rangle\}_{i=1}^d$, i.e. if there exist unique coefficients $\gamma_i \in \mathbb{C}$ such that $|\psi'\rangle = \sum_{i=1}^d \gamma_i |\psi_i\rangle$. Further, we say that the basis is *orthonormal*, if for all $i, j \in \{1 \dots d\}$, it holds that $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ where δ_{ij} is the Kronecker delta defined as

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

Hermitian and Unitary matrices Let $H \in \mathbb{C}^{d\times d}$ be a matrix. We say H is *Hermitian* if $H^\dagger = H$ where

- the symbol $\dagger$ denotes the *conjugate transpose* or *Hermitian conjugate* and acts as $(H^\dagger)_{ij} = H^*_{ji}$, while
- the symbol $*$ denotes the complex conjugate, i.e. $(a + \iota b)^* = a - \iota b$.

For a Hermitian matrix H, we say $|v\rangle$ is an *eigenvector* with *eigenvalue* v, if $H|v\rangle = \lambda|v\rangle$. We use spectrum$(H)$ to denote the set of all eigenvalues of H.

Let $U \in \mathbb{C}^{d\times d}$ be another matrix. We say U is *unitary* if $U^\dagger U = \mathbb{I}$ where

$$\mathbb{I} = \begin{bmatrix} 1 & 0 & & \\ 0 & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

is the identity matrix.

As we will state shortly, Hermitian matrices correspond to 'observables', i.e. measurable properties of the system while Unitary matrices correspond to evolution of a quantum system under a fixed Hamiltonian. They also play a pivotal role in basic linear algebra.

Exercise 16. Prove the following to ensure the notation is clear:

1. For a matrix $H \in \mathbb{C}^{d\times d}$ and a ket $|\psi\rangle \in \mathbb{C}^d$,

$$(H|\psi\rangle)^\dagger = \langle\psi| H^\dagger.$$

2. The vector

$$|\psi\rangle = \frac{|1\rangle + |2\rangle + \dots |d\rangle}{\sqrt{d}}$$

is normalised. The matrix $\Pi = \sum_{i,j=1}^{d} \frac{|i\rangle\langle j|}{d}$ can be expressed in a very simple form in terms of $|\psi\rangle$. Finally, Π satisfies $\Pi^2 = \Pi$ and has rank 1.

3. For any two orthonormal basis $\{|u_1\rangle \dots |u_d\rangle\}$ and $\{|v_1\rangle \dots |v_d\rangle\}$, there is a (unique) unitary matrix U such that $U|u_i\rangle = |v_i\rangle$ for all $i \in \{1 \dots d\}$.

We recall the following crucial fact from linear algebra, about Hermitian matrices.

Theorem 17 (Spectral Theorem)**.** *Let $O \in \mathbb{C}^{d\times d}$ be any Hermitian matrix. Then, there exists an orthonormal basis $\{|v_1\rangle, \dots |v_n\rangle\}$ such that $O = \sum_{i=1}^{d} o_i |v_i\rangle\langle v_i|$ where $o_i \in \mathbb{R}$, or in other words, there is a unitary U such that $UOU^\dagger$ is a real diagonal matrix.*

If you're seeing this for the first time, it may be worth trying to at least prove that the eigenvalues of a Hermitian matrix are real. And perhaps then come up with a proof sketch for the fact that the matrix can also be diagonalised, assuming all eigenvalues are distinct.

Exercise 18. Prove the following.

1. Every Hermitian matrix H has real eigenvalues.
2. If λ is an eigenvalue of a unitary matrix U, then $|\lambda| = 1$.
3. If λ is an eigenvalue of a projector Π (i.e. a matrix satisfying $\Pi^2 = \Pi$), then λ is either 1 or 0.

We are now in a position to write down the postulates of quantum mechanics, for a single system. We consider multiple systems later.

A quick qualification: Suppose one considers interactions between the system of interest and 'an environment', and then removes the environment. One can ask if such operations can also be included in the axioms for a single system. Indeed, they can be, and we will look at them later. For the moment, we will not consider such interactions.

Also, here we are setting all physical units such as $\hbar$ to 1 since our focus is on the information processing aspect.

With these qualifications, for a single system, quantum mechanics postulates the following.

Axiom 19 (Quantum Mechanics, for a single system)**.** *The postulates of quantum mechanics for a single system are the following.*

1. *State, Hamiltonian and Observables.*

a) ***State.*** *The state of a quantum system is specified by a normalised vector* $|\psi\rangle \in \mathcal{H}$ *satisfying* $\| |\psi\rangle \|^2 = 1$ *where* $\mathcal{H}$ *is a finite complex vector space* $\mathbb{C}^d$ *for some* $d \in \mathbb{N}$.

b) ***Observable.*** *An observable* O *of this quantum system is a Hermitian matrix acting on* $\mathcal{H}$, *i.e.* $O = O^\dagger$ *where* $O^\dagger$ *denotes entry-wise complex conjugate of* O *transpose.*

c) ***Hamiltonian****. The system's Hamiltonian is a special observable* $H(t)$ *that may depend on a real parameter* t. *This Hamiltonian determines how the state of the system changes over time, as specified in (2) below.*

2. ***Evolution.*** *Suppose the state of the system at time* t *is given by* $|\psi(t)\rangle$. *Then, the evolution of this state* $|\psi(t)\rangle$ *as a function of time is governed by the following differential equation:*

$$\frac{\partial}{\partial t} |\psi(t)\rangle = -iH(t)\, |\psi(t)\rangle \tag{2.1}$$

where $H(t)$ *is the Hamiltonian of the system.*

3. ***Measurement.*** *Suppose at time* t, *an observable* $O = \sum_i o_i \Pi_i$ *is measured, where* o_i *are the eigenvalues of* O *and* Π_i *are the corresponding projectors on the* o_i *eigenspaces.*[3] *Then, after the measurement, the system 'collapses' to the state* $|v_i\rangle$ *with probability* $\langle \psi(t)| \, \Pi_i \, |\psi(t)\rangle$ *and the observed value is* o_i.

It is an elementary exercise to verify that the expectation value of measuring an observable can be computed as follows.

Exercise 20. Verify that the expected result of measuring O when the system is in the state $|\psi\rangle$ is $\langle \psi | \, O \, |\psi\rangle$ which is sometimes briefly written as $\langle O \rangle$.

2.3.1.1 Applying these to explain Stern Gerlach

Let us see how these apply to Stern Gerlach. For now, we take $H = 0$ for simplicity and return to the $H \neq 0$ case later. For Stern Gerlach, the system (the Silver atom) is a two-level system, i.e. the corresponding vector space $\mathcal{H} = \mathbb{C}^2$.

- We assign the state $|0\rangle$ to any Silver atom that emerges on top ($+\ell\hat{z}$), even upon being repeatedly passed through SG$\hat{z}$ (recall the notation in Figure 2.2). Similarly, we assign the state $|1\rangle$ to any Silver atom that correspondingly, emerges at the bottom ($-\ell\hat{z}$) of the screen.
- We assign the observable

$$\sigma_z = |0\rangle \langle 0| - |1\rangle \langle 1| = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.2}$$

to capture the process of passing a Silver atom through SG$\hat{z}$ and marking the result as 1 if the particle lands at $+\ell\hat{z}$ and -1 if it lands at $-\ell\hat{z}$ on the screen.

Let us work out, with this convention, the result of measuring the observable σ_z when the system is in the state $|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$ where

$$|\alpha_0|^2 + |\alpha_1|^2 = 1. \tag{2.3}$$

We use the third postulate—the measurement postulate. Note that the eigenvector for eigenvalue $+1$ is $|0\rangle$ and that for -1 is $|1\rangle$. Thus, for $b \in \{0, 1\}$, the probability of seeing outcome $(-1)^b$ is given by $|\langle b|\psi\rangle|^2 = |\alpha_b|^2$, and the corresponding post-measurement state is $|b\rangle$.

Now, clearly, if one measures the post measurement state again, one will get the same result. So far, everything checks out. Now comes the crucial bit. How do we extend the definitions and the analysis to experiments involving SG$\hat{x}$? The vector space is 2 dimensional, so we cannot choose vectors independent of $|0\rangle$ and $|1\rangle$. Consider the following choice.

- We assign the state $|+\rangle := \frac{|0\rangle + |1\rangle}{\sqrt{2}}$ to any Silver atom that emerges at $+\ell\hat{x}$ and $|-\rangle := \frac{|0\rangle - |1\rangle}{\sqrt{2}}$ to any Silver atom that emerges at $-\ell\hat{x}$.
- We assign the observable

$$\sigma_x = |+\rangle \langle +| - |-\rangle \langle -| = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{2.4}$$

to capture the process of passing a Silver atom through SG$\hat{x}$ and marking the result as 1 if the particle lands at $\ell\hat{x}$ and -1 if it lands at $-\ell\hat{x}$.

[3] Non-degenerate means that all the eigenvalues are distinct.

For simplicity, let us suppose that the state of a Silver atom out of the oven is given by $|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$ where α_0 and α_1 are random complex numbers, satisfying Equation (2.3).

With these conventions, it is not hard to see that the results of all three experiments in Definition 7 coincide with those stated in Fact 8. Using our formalism, one can write down the three experiments as follows:

1. (a) Start with a Silver Atom in the state$|\psi\rangle$ as described above
   (b) Measure σ_z and proceed only when $+1$ outcome is obtained
   (d) Measure σ_z and report the outcome
2. (a) Start with a Silver Atom in the state$|\psi\rangle$ as described above
   (b) Measure σ_z and proceed only when $+1$ outcome is obtained
   (d) Measure σ_x and report the outcome
3. (a) Start with a Silver Atom in the state$|\psi\rangle$ as described above
   (b) Measure σ_z and proceed only when $+1$ outcome is obtained
   (c) Measure σ_x and proceed only when $+1$ is obtained
   (d) Measure σ_z and report the outcome

We can now see quantitatively what role the extra measurement σ_x plays, to produce the difference between experiments 1 and 3. In both 1 and 3, right after step (b), the Silver Atom is in the state $|0\rangle$. If σ_z is measured again, the result is $+1$ with probability 1, and the post measurement state remains unchanged. However, if σ_x is measured, one obtains $|+\rangle$ with probability $1/2$ and $|-\rangle$ with probability $1/2$. After *post-selecting* on $|+\rangle$ (i.e. proceeding as in Experiment 3), when σ_z is measured, the probability of obtaining outcome $+1$ is $1/2$ and that of -1 is also $1/2$ (because $|\langle +|0\rangle|^2 = |\langle -|0\rangle|^2 = 1/2$).

Mathematically, the difference in Experiment 1's output and Experiment 2's output is arising from the fact that σ_z and σ_x do not *commute*. More precisely, we have the following.

Notation 21. The following is now standard notation.

- *Pauli Matrices.* The matrices σ_x, σ_y and σ_z are termed *Pauli Matrices* where σ_z and σ_x are as in Equation (2.2) and (2.4) resp. while

$$\sigma_y := |\tilde{+}\rangle\langle\tilde{+}| - |\tilde{-}\rangle\langle\tilde{-}| = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{2.5}$$

  for $|\tilde{\pm}\rangle := (|0\rangle \pm i\,|1\rangle)/\sqrt{2}$.
- *Commutator.* For any two operators, A and B, the *commutator* is defined as $[A, B] := AB - BA$.

Exercise 22. Let σ_x, σ_y and σ_z be as above (i.e. Equation (2.4), (2.5) and (2.2) respectively). Then, show that

$$\begin{aligned} [\sigma_z, \sigma_x] &= 2i\sigma_y \\ [\sigma_y, \sigma_z] &= 2i\sigma_x \\ [\sigma_x, \sigma_y] &= 2i\sigma_z \end{aligned}$$

Also, verify that for all $\tau \in \{\sigma_x, \sigma_y, \sigma_z\}$,

1. $\tau^2 = \mathbb{I}$ and so $[\tau, \tau] = 0$,
2. $\det(\tau) = -1$ and $\mathrm{tr}(\tau) = 0$.

While σ_z and σ_x made at least some sense relative to the Stern Gerlach experiment, σ_y seems a bit more mysterious. But one can imagine, a frame of reference where the magnetic field is inhomegenous along $\hat{y}$ and the Silver atoms have a velocity solely along say $\hat{z}$. Then, SG$\hat{y}$ makes sense and one can then ascribe the state $|\tilde{+}\rangle$ to states landing at $+\ell\hat{y}$ and $|\tilde{-}\rangle$ to those at $-\ell\hat{y}$. This shows how the Pauli Matrices are related to physical orientation.

More formally, one can verify that the matrices $i\sigma_x, i\sigma_y, i\sigma_z$ form a basis for Lie algebra $\mathfrak{su}(2)$ which exponentiates to give the special unitary group $\mathrm{SU}(2)$. The group $\mathrm{SU}(2)$ is a double cover of the group of rotations in three dimensions $\mathrm{SO}(3)$.

Theorem 23. *The Lie group of rotations in three dimensions,* $\mathrm{SO}(3)$ *is isomorphic to that of the special unitary group,* $\mathrm{SU}(2)$*, i.e.* $(\mathbb{R}^3, \times) \cong \mathfrak{so}(3) \cong \mathfrak{su}(2)$.

All this to say that there is a strong connection rotations in three dimension and the Pauli matrices we introduced. We will not need the details or proofs beyond this for cryptographic purposes.

We end this discussion by introducing the Bloch sphere, which gives a nice geometric interpretation (in three dimensions) to the vectors we introduced to describe the Silver atom.

Notation 24 (Bloch Sphere). Consider the following parametrisation: $|\theta,\phi\rangle := \cos\theta/2\,|0\rangle + e^{i\phi}\sin\theta/2\,|1\rangle$ where $\theta \in [0,\pi]$ and $\phi \in [0, 2\pi)$. This allows one to interpret the vector $|\theta,\phi\rangle$ as a vector on the surface of a three dimensional sphere by taking θ as the polar angle and ϕ as the azimuthal, as depicted in Figure 2.4.

Note 25. Observe that the eigenvectors of σ_z, σ_x, and σ_y (resp.) on the three dimensional sphere, correspond to vectors along the z, x and y axis, as detailed below

$$\begin{aligned}
|0\rangle &= |0,\phi\rangle && \text{for any } \phi \\
|1\rangle &= |\pi,\phi\rangle && \text{for any } \phi \\
|+\rangle &= |\pi/2, 0\rangle \\
|-\rangle &= |\pi/2, \pi\rangle \\
\left|\tilde{+}\right\rangle &= |\pi/2, \pi/2\rangle \\
\left|\tilde{+}\right\rangle &= |\pi/2, 3\pi/2\rangle
\end{aligned}$$

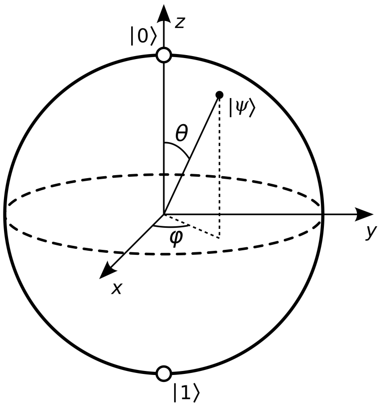

Figure 2.4: Bloch Sphere (taken from Wikipedia)

2.3.1.2 The uncertainty principle

Note that if two observables O_1 and O_2 commute, i.e. $[O_1, O_2] = 0$, then the order in which they are measured, does not matter. This is a direct consequence of the following fact.

Fact 26. *Let O_1, O_2 be Hermitian matrices in $\mathbb{C}^d$. If $[O_1, O_2] = 0$, then there is a basis $\{|v_1\rangle, \dots |v_d\rangle\}$ in which O_1, O_2 are simultaneously diagonal, i.e. $O_1 = \sum_i \lambda_i^{(1)} |v_i\rangle\langle v_i|$ and $O_2 = \sum_i \lambda_i^{(2)} |v_i\rangle\langle v_i|$.*

Proof sketch. The key observation to understand this fact is the following: Consider an eigenvector $|v\rangle$ of O_1 with eigenvalue λ. Then, $O_2|v\rangle$ is also an eigenvector of O_1 with eigenvalue λ since

$$O_1(O_2|v\rangle) = O_2 O_1 |v\rangle = \lambda(O_2|v\rangle) \tag{2.6}$$

. How does this help? To illustrate the point, assume that O_1 has a two eigenvectors with eigenvalue λ (if it had 1, then there is nothing to do; and the more than two case follows analogously).

Now, given this assumption, Equation (2.6) shows that O_2 leaves O_1's λ-valued eigenspace invariant. Thus, one can restrict O_2 to this eigenspace, i.e. $\mathsf{span}\{|v_1\rangle, O_2|v_1\rangle\}$, and diagonalise O_2. The resulting basis will be an eigenbasis for both O_1 and O_2, restricted to this subspace. The argument can be repeated for all λ in the spectrum of O_1. □

Using Fact 26, one can show that indeed the measurement order does not matter when the observables commute.

Exercise 27. Let O_1, O_2 be as above and let $|\psi\rangle \in \mathbb{C}^d$ denote any quantum state. Consider the following experiments:

1. Measure O_1 first, to obtain o_1 and then measure O_2 to obtain o_2
2. Measure O_2 first, to obtain o_2 and then measure O_1 to obtain o_1