1

Chapter

Introduction

Starting 2018, an incredible amount of progress has been made in this burgeoning area called quantum cryptography. In these notes, my goal is to give you some flavour of these exciting developments. The following is an ambitious and therefore tentative course outline. The basic goal will be to cover Unit 1 and one key result from Units 2, 3 and 4. The remaining units will be covered, depending on the pace of the course.

§ 1.1 Outline

Unit 1: Review:

Primer on quantum formalism. [VW '16]

Lay of the land:

Impagliazzo's worlds: in particular MiniCrypt and Cryptomania

Introduce the main directions:

- (T1) post-quantum cryptography (make classical constructions secure against quantum adversaries),
- (T2) quantum analogues of classical functionalities,
- (T3) functionalities impossible without quantum (excluding those in T2),
- (T4) basing cryptography on quantum complexity

Focus on information theoretic results (T3): key distribution (BB84, Ekert), proof of security, secret sharing, impossibility of bit commitment, impossibility of strong coin flipping, achieving optimal strong coin flipping using weak coin flipping, self-testing CHSH, all bipartite states can be self-tested [VW '24].

Unit 2: (T2) Verification:

Regev's quantum reduction [Regev'09]

Weak quantum verifier: based on MBQC [GKK17]

Classical verifier: assuming LWE is hard, Mahadev [Mahadev 18, Vidick 22]

Classical verifier: two non-communicating provers [RUV 12]

Unit 3: (T2) QFHE: construction and its applications:

Mahadev's QFHE construction [Mahadev 17]

Compiling non-local games [KLVY 23]

Verification assuming QFHE [NZ 23]

Unit 4: (T4) Minimal assumptions: Minicrypt and below

OT is in Miniqcrypt [GLSV 20, BCKM 20]

Crypto despite having NP=P or similar [Kre 21, KQST 23]

Microcrypt Primitive Zoo [Or Sattah]

Unit 5: (T1/T2) Multi Party Computation

(T1) Post-quantum Commitments (collapse-binding commitments) [Unruh'16]

(T1) Post-quantum MPC [HSS 11]

(T2) Quantum 2-PC [DNS '10 and '12]

(T2) Quantum MPC w/ quantum communication [Dulek, Grilo, Jeffery, Majenz, Schaffner]

(T2) Quantum MPC w/ classical communication [Bartusek '21]

Unit 6: (T3) Quantum-only functionalities I

Unclonable encryption: construction in the Random Oracle Model [AKLLZ '22]

Certified deletion [BK '22]

Quantum Pseudorandom unitaries [MH '24]

Unit 7: (T3) Quantum-only functionalities II

Quantum Money and Lightning [Zhandry '17]

iO for pseudo-deterministic functions [BKNY '23]

Unit 8: Bonus/extra reading (references will be provided later)

Other key topics in cryptography:

Interactive proofs

Zero knowledge

Quantum rewinding

Quantum Random Oracle Model

Everlasting security

Connections to physics:

Self-testing using a single quantum device

Non-locality => Proof of quantumness + rigidity

Black hole radiation decoding and commitments

Cryptographic tests of python's lunch conjecture

Computationally bounded theory of entanglement

References

GKK 17: Verification of quantum computation: An overview of existing approaches

VW 24: Introduction to Quantum Cryptography (book)

VW 16: Quantum Proofs (survey)

RUV 12: Classical command of quantum systems

Vidick 22: Course FSMP, Fall'20: Interactions with Quantum Devices

Mahadev 18: Classical verification of quantum computations

Mahadev 17: Classical Homomorphic Encryption for Quantum Circuits

KLVY 23: Quantum Advantage from Any Non-Local Game

NZ 23: Bounding the Quantum Value of Compiled Nonlocal Games: From CHSH to BQP verification

BCKM: One-Way Functions imply Secure Computation in a Quantum World

MH24: How to Construct Random Unitaries

Or Sattah: MicroCrypt Zoo

Kre21: Quantum pseudorandomness and classical complexity

KQST21: Quantum cryptography in algorithmica

DNS '10: Secure two-party quantum evaluation of unitaries against specious adversaries

DNS '12: Actively Secure Two-Party Evaluation of any Quantum Operation

DGJMS '19: Secure Multi-party Quantum Computation with a dishonest majority

Bartusek '21: Secure Quantum Computation with Classical Communication

BK 22: Cryptography with Certified Deletion

HSS11: Classical Cryptographic Protocols in a Quantum World

Unruh'16: Collapse-binding quantum commitments w/o random oracles

AKLLZ '22: On the Feasibility of Unclonable Encryption and more

Zhandry '17: Quantum Lightning Never Strikes the same state twice

MH '24: How to Construct Random Unitaries

BKNY '23: Obfuscation of pseudo-deterministic quantum circuits

Regev '09: On Lattices, Learning with Errors, Random Linear Codes, and Cryptography

Video Tutorials:

- Dakshita Khurana: Cryptography with certified deletion
- Mark Zhandry: Security Reductions (multi-part series)

Preferred/reference Textbooks:

There are no textbooks for this course as the material is at the frontier of current research in quantum cryptography. Relevant lecture notes/tutorials/survey articles have already been referenced above.

Resources to review basics of quantum info/computation

- Lecture notes on Quantum Computation by John Preskill (Caltech)
- Introduction to Quantum Information and Computation, MA Nielsen and IL Chuang

Resources to review classical cryptography

- Introduction to Modern Cryptography. Jonathan Katz and Yehuda Lindell.
- Foundations of Cryptography (Volumes 1 and 2). Oded Goldreich.

Other resources

- Zhandry's lecture notes on quantum cryptography

2

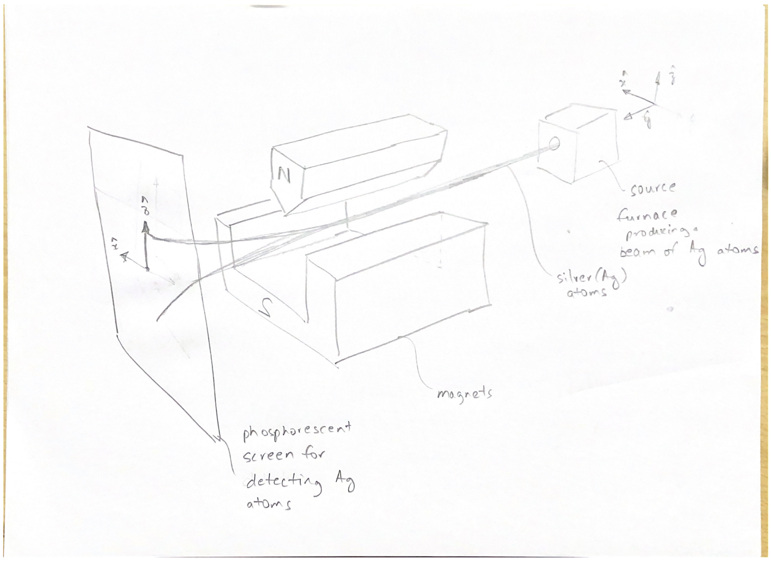
Chapter

Quantum Mechanics

This is a condensed and simplified survey of the key concepts and results in quantum mechanics. While it does not assume any prior knowledge of quantum mechanics, readers are highly encouraged to study introductory material (such as [Neilsen Chuang], Preskill's notes) on the subject to get a solid understanding.

§ 2.1 Stern-Gerlach

Before we describe the quantum formalism, which can seem quite abstract, let us first look at a physics experiment that conceptually (not historically) shows why physicists were forced to depart from classical physics.



$\hat{x}$
$\hat{z}$
$\hat{y}$
N
S
$\hat{z}$
$\hat{x}$
source
furnace producing a beam of Ag atoms
silver (Ag) atoms
magnets
phosphorescent screen for detecting Ag atoms

Figure 2.1:

Definition 1 (Stern-Gerlach)**.** By a *Stern-Gerlach experiment*, we mean a setup that is illustrated in Figure 2.1. Here, silver atoms are heated in a furnace and let out through a single opening. These are channeled into a beam and passed through a magnetic field gradient. The magnetic field gradient (conventionally taken along the z axis) is produced by having two magnets with different shapes—one wedge shaped and the other flat—as shown. Finally, the beam is collected on a phosphorescent screen which serves as a detector for the silver atoms.

All this might look a bit arbitrary at first. Why Silver atoms? Why the magnetic field gradient? Why the furnace? But there is good reason for each choice.

Since we only have so much time, we will have to take some results from physics as facts.

Fact 2. *Consider a point particle, that is electrically neutral. Suppose this particle has a 'magnetic moment' μ_z along the z-axis. If this particle is placed in a magnetic field, then the force F_z experienced by the particle*[1] *is given by*

$$F_z = \mu_z \frac{\partial B_z}{\partial z}$$

[1] In general, the force is given by $\mathbf{F} = \nabla(\mu \cdot \mathbf{B})$

where B_z is the z component of the magnetic field.

Above, we did not define the term 'magnetic moment' but we do not need to go into the details here. An intuitive understanding suffices. For instance, a loop carrying current I has magnetic moment IS where S is the surface area of the loop. The 'magnetic moment' determines how strongly a substance interacts with a magnetic field.

We know from other experiments that electrons have a kind of 'intrinsic spin'. Since they are charged, this spin produces a magnetic moment, analogously to how a loop carrying current produces a magnetic moment. This analogy breaks down on closer inspection—the intrinsic spin of an electron turns out to behave quite differently from that produced by a current carrying loop. To uncover this behaviour, Silver atoms turn out to be perfect candidates.

Fact 3. *A Silver atom,* Ag*, has 47 electrons (and it is electrically neutral). Its magnetic moment μ is proportional to the 'intrinsic spin' S of its outermost electron, i.e.*

$$\mu = \frac{e}{m_e c} S$$

where e (is negative) is the charge of the electron, m_e its mass and c the speed of light.

The Silver atom has the nice property that its magnetic moment is essentially the same as that produced by its outermost electron, allowing us to isolate and better understand this mysterious notion of an intrinsic spin of an electron.

Question 4. *Suppose the Stern-Gerlach experiment (as in Definition 1) is executed as described. What pattern should someone who does not know anything about quantum mechanics, expect to see on the screen?*

Let us treat the spin S of a Silver atom classically, i.e. as a vector, denoting the angular momentum of the atom. Suppose that the magnitude of S is a fixed property of the atom. With these simplifying assumptions, we can answer Definition 1.

Suppose S points upwards (i.e. along the unit vector $\hat{z}$ along the z-axis). Suppose the atom lands at position $\ell\hat{z}$ where ℓ is determined by the experimental setup. By symmetry then, it is clear that when S points downwards, the atom should land at $-\ell\hat{z}$. Now, if S pointed along the direction of the beam, then $\mu_z = 0$ and hence $F_z = 0$. Thus, in this case the atom arrives at the centre, i.e. at $0\hat{z}$. In the experiment, since the atoms are being released from a furnace, it is reasonable to assume that the direction of spin of the atoms in the beam are distributed uniformly randomly. Thus, for each position z between $-\ell$ and ℓ along the z-axis, there the probability density $p(z)$ for seeing a Silver atom on the screen, is non-zero. One can also compute $p(z)$ with some elementary calculations. However, the key observation is that the distribution is continuous because in reality, when this experiment is performed, the following is observed.

Fact 5. *After execution of the Stern-Gerlach experiment (as in Definition 1), the screen shows deposits of Silver atoms in two spots: one located at ℓ meters above the centre, and one below, the centre (where ℓ, as discussed above, is determined by the experimental setup and by centre, we mean the point where the beam of Silver atoms would appear had there been no magnets).*

The spin seems to be 'quantised', it either points up or down. Clearly, our Silver atoms are defying our classical intuition. How might we explain this observed behaviour? Given just this observation, it is reasonable posit the following:

Assumption 6. *Each Silver atom can have only one of two values of spin, along any component, i.e. S_z, S_y, S_x each takes a (possibly distinct) value in ± 1 (in some appropriate units). (Silver atoms emerging from the furnace, are equally likely to have any of these possible values of spin).*

If this is indeed how Silver atoms behave, then the result of the Stern-Gerlach experiment can be explained. Or can it? Consider the following sequential Stern-Gerlach experiment that are designed to test this assumption.

Definition 7 (Sequential Stern Gerlach Experiments)**.** Let SG$\hat{z}$ denote the magnets of the Stern Gerlach setup, as in Definition 1 and let SG$\hat{x}$ denote the same but with the magnets rotated to produce an inhomogeneous magnetic field along the x-axis instead of the z-axis.

In Definition 1, denote the beam corresponding to the particles landing at $\ell\hat{z}$ by S_{z+} and those corresponding to $-\ell\hat{z}$ by S_{z-}. When the setup is given SG$\hat{x}$ (instead of SG$\hat{z}$) denote the beam by S_{x+} and S_{x-} respectively.

We can now specify the following three, sequential Stern-Gerlach experiments (as shown in Figure 2.2).

1. Silver atoms come out of the oven, are subject to SG$\hat{z}$, the S_{z-} beam is blocked and the S_{z+} beam is subject to another SG$\hat{z}$ setup.

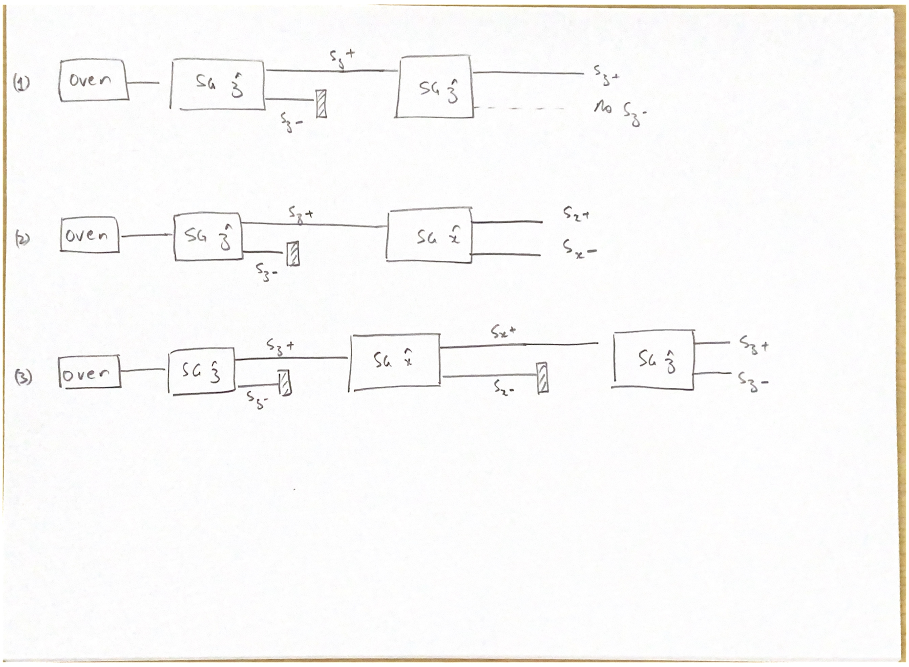


Figure 2.2: Sequential Stern Gerlach experiments (see Definition 7)

2. Same as above, except the second SG$\hat{z}$ is replaced with SG$\hat{x}$.
3. The setup in (2) is executed but the beams are not detected. Instead, the S_{x-} beam is blocked, and the S_{x+} beam is subject to a SG$\hat{z}$ again.

It is not hard to work out the output of running the experiments as stated in Definition 7, taking Assumption 6 to be true.

In experiment (1), after the first SG$\hat{z}$ the Silver atoms with a negative S_z component are removed and so when the S_{z+} beam is passed through the second SG apparatus, only the S_{z+} beam emerges—no S_{z-} should appear.

In experiment (2), after the first SG$\hat{z}$ the Silver atoms with a negative S_z component are removed. However, the S_{z+} beam contains Silver atoms with both ± 1 spin components along S_x. Hence, when the S_{z+} beam is passed through SG$\hat{x}$, we expect both the S_{x+} and the S_{x-} beam to be present.

And indeed, these predictions checkout in the lab. However, the third experiment is where Assumption 6 becomes untenable.

In experiment (3), according to Assumption 6, after SG$\hat{x}$, both the S_{x+} and S_{x-} beam must only have the positive component of S_z. Hence, when S_{x+} is passed through SG$\hat{z}$, we do not expect an S_{z-} beam. But in the lab, both the S_{z+} and S_{z-} beams are observed!

Fact 8. *Suppose that Experiments (1), (2) and (3) from Definition 7 are performed. Then, their respective results are the following:*

1. *Only beam S_{z+} is detected*
2. *Both beams S_{x+} and S_{x-} are detected*
3. *Both beams S_{z+} and S_{z-} are detected*

These observations tell us that Nature is much more delicate and interesting. Attempts such as those based on Assumption 6 are inadequate. Physicists realised that, based on experiments like such as this one, that a radical departure from 'classical physics' was needed to explain the workings of Nature at deeper levels. This effort has been a remarkable success and is now termed quantum physics.

We end our discussion here by noting two peculiar features of quantum mechanics, anticipating the axiomatic treatment in a later section. First, note that almost every physical property in our experience, tends to be continuous—position, momentum, angular momentum, potential energy, kinetic energy, temperature etc. And yet, the Stern-Gerlach experiment reveals that at a fundamental level, Nature is *quantised.* Second, measuring the x component of angular momentum, disturbs the z component of the angular momentum. In more detail, Experiment (1) in Definition 7, may be interpreted as first preparing particles with the z component of angular momentum pointing upwards, to produce the S_{z+} beam, and then measuring the z component of angular momentum using the second SG$\hat{z}$ apparatus. Experiment (3) then, does the same, except that it uses SG$\hat{x}$ in between (the preparation and z component measurement), to measure the x component of angular momentum. And this changes the result of the experiment. This, as we will formalise shortly, shows that not all measurements are compatible with each other, i.e. the *order of measurements matters.*

Exercise 9. Look up (for instance in Sakurai's book) how polarised light also exhibits the second feature discussed above.

§ 2.2 The double slit experiment

Another experiment that is instructive for any student of quantum mechanics, is the double slit experiment. It was originally conceived by Thomas Young in 1801 to demonstrate the wave nature of light. It was later found that electrons and even atoms, show the same behaviour! And, when this experiment is performed using individual particles (i.e. one particle at a time), it demonstrates that measurements play a central role in our current quantum formalism.

The experimental setup is strikingly simple.

Definition 10 (Young's Double-Slit Experiment)**.** The double slit experiment (see Figure 2.3) consists of a source that produces a beam of light normal to an opaque plate with two parallel slits. A screen is placed behind this plate to view the pattern.

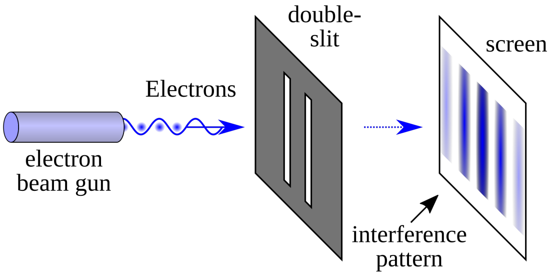


Figure 2.3: Young's Double-Slit experiment (taken from https://commons.wikimedia.org/wiki/File:Double-slit.PNG).

This initial experiment was conceived to test whether light behaves as a collection of particles or as a wave. Anyone who has thrown pebbles in water is well aware of the interference patterns that waves naturally generate. With some elementary calculation, one can check that indeed, using the electromagnetic wave equation for light, one obtains an interference pattern on the screen, parallel to the slits, as shown in Figure 2.3. On the other hand, if light consisted of particles, one would expect two peaks in the pattern on the screen: one corresponding to each slit. While connection between electromagnetism and light was not known at the time, this experiment convincingly settled the question of whether light behaves as a wave or a collection of particles.

Fact 11. *An interference pattern as shown in Figure 2.3 is obtained, when the experiment in Definition 10 is conducted.*

However, as it turns out, Nature is more subtle and beautiful. In his 1924 PhD thesis, Louis de Broglie postulated that all matter has wave properties. This was confirmed by essentially the same double slit experiment, in 1927, where instead of light, electrons were used.

Fact 12. *In the experiment stated in Definition 10, when light is replaced by matter (such as electron, atoms or even molecules), the conclusions of Fact 11 hold.*

For light, Young's experiment makes sense because we now know light is just an electromagnetic wave. However, for matter, what does this wave describe? Even besides issues with interpretation, there is a more glaring operational concern: what happens if one repeats this experiment with a single electron. Surely, a single electron has to pass through one of the slits, right?

Let us try to answer the second question as it will help us answer the first question. Consider the following two variants of the double slit experiment.

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

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

and the corresponding *bra* $\langle\psi|$ in $\mathcal{H}$ is written as

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

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

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

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

Basis and orthonormal vectors We say 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 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 † denotes the *conjugate transpose* or *Hermitian conjugate* and acts as $(H^\dagger)_{ij} = H_{ji}^*$, while
- the symbol $*$ denotes the complex conjugate, i.e. $(a + \imath b)^* = a - \imath b$.

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 it later. The system, the Silver atom, in that case 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 $\text{tr}(\tau) = 0$.

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: $|\theta,\phi\rangle := \cos 2\theta\,|0\rangle + e^{i\phi}\sin 2\theta\,|1\rangle$ 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

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

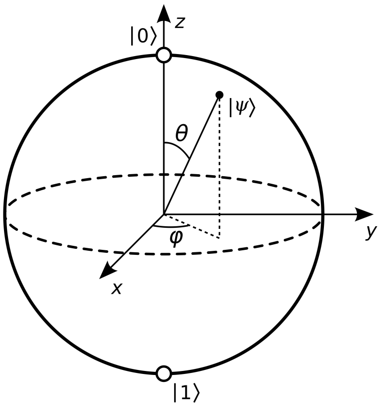


Figure 2.4: Bloch Sphere

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 $\{|v_1\rangle, \ldots |v_d\rangle\}$ 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 $O_1(O_2|v\rangle) = O_2O_1|v\rangle = \lambda(O_2|v\rangle)$. How does this help? It means that O_2 leaves O_1's λ-valued eigenspace invariant. Thus, one can restrict O_2 to this eigenspace, 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

Prove that the probability one obtains $o_1 = a$ and $o_2 = b$ is the same in both cases, i.e.

$$\Pr[(o_1, o_2) = (a, b) : \text{Experiment 1}] = \Pr[(o_1, o_2) = (a, b) : \text{Experiment 2}]$$

for all $a \in \text{spectrum}(O_1)$ and all $b \in \text{spectrum}(O_2)$.

The discussion around commuting observables is interesting because it shows that in quantum systems, some properties cannot be simultaneously measured with arbitrary precision. For instance, measuring the x component of the spin of an electron, i.e. measuring σ_x, will 'disturb' the spin of the electron along the other components. This is not a limitation of the measurement apparatus but a fundamental limitation imposed by quantum theory and the fact that $[\sigma_x, \sigma_y] \neq 0$ and $[\sigma_x, \sigma_z] \neq 0$. While here we are not dealing with continuous variables, it is worth mentioning that with position x and momentum p of a particle, essentially the same issue arises. The limit on precisely measuring both position and momentum was famously quantified by Heisenberg as follows.

For any observable O, and quantum state $|\psi\rangle$, let $\Delta O := O - \langle O \rangle$ where, recall, $\langle O \rangle = \langle \psi | O | \psi \rangle$. Then, the expectation value of $(\Delta O)^2$ captures the 'dispersion' in the observed value of O from the expected value of O. To see this, note that if $|\psi\rangle$ is an eigenvector of O, then $\left\langle (\Delta O)^2 \right\rangle = 0$.

Theorem 28 (Uncertainty principle)**.** *Let O_1, O_2 be two observables. Then for every state $|\psi\rangle$, the uncertainty as quantified by the dispersion in O_1 and O_2 together, is lower bounded by the commutator of O_1 and O_2, i.e.*

$$\left\langle (\Delta O_1)^2 \right\rangle \left\langle (\Delta O_2)^2 \right\rangle \geq \frac{1}{4} \left| \langle [O_1, O_2] \rangle \right|^2 .$$

We will prove variants of this statement as needed later. For now, we omit the proof and note that the key ingredient in the proof is the Cauchy-Schwarz inequality.

Lemma 29 (Cauchy-Schwarz inequality)**.** *For two vectors $|u\rangle, |v\rangle \in \mathbb{C}^d$, it holds that $\langle u|u\rangle \langle v|v\rangle \geq |\langle u|v\rangle|^2$.*

2.3.1.3 Applying the axioms to the double slit experiment

I wanted to include an explanation of the double slit experiment using Axiom 19 but in the interest of time, I leave this as an exercise.

Exercise 30. Understand what the Mach-Zhander interferometer is. Can one view it as a simplification of the Double-Slit experiment (see Definition 10 and (13))? Explain the observations made using a Mach-Zhander interferometer starting from Axiom 19.

2.3.1.4 Solving the Schrödinger Equation

So far, we took $H = 0$. We now see what happens when $H \neq 0$. Recall that H only appears in the differential equation, Equation (2.1), Postulate 2. For our purposes, this can be simplified to the point that we do not need to worry about differential equations. To see this, we first recall the following fact.

Fact 31 (Hermitian Exponentiation)**.** *Consider a vector space $\mathbb{C}^d$ and denote by $\mathrm{U}(d)$ the set of all unitary matrices, i.e.*

$$\mathrm{U}(d) := \{U : U^\dagger U = \mathbb{I}\}.$$

Then it holds that

$$\mathrm{U}(d) = \{e^{iH} : H = H^\dagger\}.$$

When H is restricted to also have zero trace, then the group becomes the special unitary group, i.e.

$$\mathrm{SU}(d) = \{e^{iH} : H = H^\dagger, \mathrm{tr}(H) = 0\}$$

where $\mathrm{SU}(d) := \{U : U^\dagger U = 1, \det(U) = 1\}$.

The second relevant observation is the following.

Claim 32 (Solution to Schrödinger's equation for time independent Hamiltonians). Using the notation in Axiom 19, suppose $H(t)$ is independent of time. Then the solution to Equation (2.1) is

$$|\psi(t')\rangle = e^{-i(t'-t)H} |\psi(t)\rangle$$

for all t', t.

Using the claim above and Fact 31, it is clear that by choosing H appropriately, one can apply any unitary U to $|\psi\rangle$.

However, not all unitaries are 'easy' to implement so there is more to be said. We cover this later when we discuss the circuit model in Subsection 2.4.1.

2.3.2 Multiple systems

These axioms were for a single system. Suppose two systems are involved, such as two different Silver atoms. How does quantum mechanics apply to these? To this end, we introduce some notation anticipating its use in computing and cryptographic contexts. We are following the convention in [2].

Notation 33. We use the word *register* to refer to a physical system (such as a Silver atom). We typically use capital letters in sans serif font such as $\mathsf{X}, \mathsf{Y}, \mathsf{Z}$ to refer to registers and calligraphic letters $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ to refer to their corresponding Hilbert spaces.

Definition 34 (Quantum state of multiple registers/systems)**.** Given two registers X and Y, the combined state of the two registers is given by vector in $\mathcal{X} \otimes \mathcal{Y}$ where the vector space $\mathcal{X} \otimes \mathcal{Y}$ is spanned by $\{|u\rangle \otimes |v\rangle : |u\rangle \in \mathcal{X}, |v\rangle \in \mathcal{Y}\}$. We also use $|uv\rangle$ or $|u\rangle\,|v\rangle$ to write $|u\rangle \otimes |v\rangle$ briefly.

Informally, for those not familiar with the tensor product notation, when we write $|u\rangle \otimes |v\rangle$, we basically need to keep two properties of the tensor product in mind

1. $(|u\rangle \otimes |v\rangle)^\dagger = \langle u| \otimes \langle v|$ and
2. $(\langle u'| \otimes \langle v'|)\,(|u\rangle \otimes |v\rangle) = \langle u'|u\rangle\,\langle v'|v\rangle$.

With these, it is easy to check the following.

Exercise 35. If $\{|x_i\rangle\}_{i \in I}$ and $\{|y_j\rangle\}_{j \in J}$ are orthonormal bases for vector spaces $\mathcal{X}$ and $\mathcal{Y}$ respectively, then $\{|x_i\rangle \otimes |y_j\rangle\}_{i \in I, j \in J}$ spans the vector space $\mathcal{X} \otimes \mathcal{Y}$.

To see why the tensor product appears, perhaps it is worth looking at how similar situations are handled in classical probability theory. Consider a process that involves rolling two die. Now each dice can produce 6 outcomes. However, when two die are concerned, the sample space of outcomes becomes 6×6. Intuitively, quantum theory is generalising this idea to vector spaces, using tensor products.

Now that we understand how to describe the quantum state of multiple registers/systems, it is easy to extend Axiom 19—they remain unchanged, except that the initial state must be taken to be a vector in the tensor product of the Hilbert spaces corresponding to the various systems.

What is more interesting now, is that one can ask the following question: Suppose one has two registers S and E, denoting the system of interest and the environment, in the state $|\psi\rangle_{\mathsf{S}}\,|\phi\rangle_{\mathsf{E}}$. Suppose S and E are made to interact by applying some Hamiltonian H_{SE} acting on both $\mathcal{S}$ and $\mathcal{E}$ (the corresponding Hilbert spaces of registers S and E). And finally, system E is removed. Given H_{SE} and $|\phi\rangle$, clearly, the system S goes from the state $|\psi\rangle$ to some different $\mathcal{C}(|\psi\rangle)$. Clearly, this is a physical process and yet, one can show that the transformation from $|\psi\rangle$ to $\mathcal{C}(|\psi\rangle)$ is not unitary in general. To see this, we first introduce the notion of mixed and pure states. And then, we state an equivalent version of the axioms of quantum mechanics to allow for such transformations, without having to explicitly invoke the existence of an environment.

2.3.2.1 Pure and Mixed States (tracing out systems)

So far, we have only discussed what are called *pure states*—the state of a system S is a vector $|\psi\rangle \in \mathcal{S}$ in the Hilbert space corresponding to S. However, it is easy to see that the formalism we have introduced does not handle classical probabilities conveniently. Suppose that Aman prepares system S in the state $|\psi_0\rangle$ and $|\psi_1\rangle$ with probability p_0 and p_1 respectively[4] and gives this state to Basanti, without telling her which state he prepared. How should Basanti write the state of system S, from her point of view?

First, note that the axioms of quantum mechanics (as in Axiom 19) allow Basanti to compute what happens in any given experiment. For instance, suppose Basanti evolves the system using the unitary U and then measures the observable O. Then, the expected value of the observable, can be computed by computing what happens in each case—when the state is $|\psi_0\rangle$ and when the state is $|\psi_1\rangle$—and then weighting these outcomes with p_0 and p_1. More precisely, the expected value O will be

$$p_0 \langle\psi_0|\, U^\dagger O U \,|\psi_0\rangle + p_1 \langle\psi_1|\, U^\dagger O U \,|\psi_1\rangle\,. \tag{2.6}$$

Such calculations can be done for simple situations but become quite cumbersome in general.

To remedy this, the *density matrix formalism* was introduced, which allows one to handle both pure states and mixed states—classical mixtures of pure states states (like the one discussed above).

[4] We impose $p_0 < 1$ and $p_1 = 1 - p_0$ to ensure these really are probabilities.

Definition 36. Let S be a register and $\mathcal{S} = \mathbb{C}^d$ be its associated vector space. Then the state of S is given by a *density matrix* ρ in $\mathbb{C}^{d\times d}$ which satisfies the following conditions:

1. ρ is Hermitian,
2. $\rho \geq 0$, i.e. all its eigenvalues are non-negative
3. $\mathsf{tr}(\rho) = 1$.

The set of all density matrices is denoted by $\mathsf{D}(\mathcal{S})$.

Let us write down Basanti's state using this density matrix formalism:

$$\rho = p_0 |\psi_0\rangle \langle\psi_0| + p_1 |\psi_1\rangle \langle\psi_1| .$$

Note that one can write $\langle\psi_0| O |\psi_0\rangle = \mathsf{tr}(O |\psi_0\rangle \langle\psi_0|)$ and therefore the expected value of O can be computed as $\mathsf{tr}(O\rho)$ which matches the expression in Equation (2.6). It is straightforward to observe how Axiom 19 applies to density matrices. Since ρ is a Hermitian matrix, it can always be diagonalised (see Theorem 17). Writing $\rho = \sum_{i=1}^{d} p_i |\phi_i\rangle \langle\phi_i|$ where $p_i \geq 0$ and $\{|\phi_i\rangle\}_{i=1}^{d}$ are orthonormal, we note that if the states $|\phi_i\rangle$ are mapped to $|\phi'_i\rangle$ using Axiom 19, then the corresponding density matrix $\rho' = \sum_{i=1}^{i} p_i |\phi'_i\rangle \langle\phi'_i|$. For instance, if $|\phi'\rangle = |\phi(t)\rangle = e^{-itH} |\phi\rangle = U |\phi\rangle$, ρ maps to $U\rho U^\dagger$.

Exercise 37. Prove that for any state $|\psi\rangle \in \mathbb{C}^d$ and any matrix $O \in \mathbb{C}^{d\times d}$, $\langle\psi| O |\psi\rangle = \mathsf{tr}(O |\psi\rangle \langle\psi|)$.

Before we return to the question of interacting a system with an environment and then removing the environment, we briefly mention a few noteworthy properties of density matrices. First, even though we motivated the density matrix by saying that a classical mixture of states $|\psi_0\rangle$ and $|\psi_1\rangle$ can be viewed as a density matrix, the inverse is not true—given a density matrix, (in general) there is no unique set of pure states and probabilities that constitute the density matrix.

Exercise 38. Let $\rho := \frac{3}{4} |0\rangle \langle 0| + \frac{1}{4} |1\rangle \langle 1|$ and $\sigma := \frac{1}{2} |u\rangle \langle u| + \frac{1}{2} |v\rangle \langle v|$ where $|u\rangle := \sqrt{3/4}\, |0\rangle + \sqrt{1/4}\, |1\rangle$ and $|v\rangle := \sqrt{3/4}\, |0\rangle - \sqrt{1/4}\, |1\rangle$. Is $\rho = \sigma$?

The other property is the following observation which follows immediately from Theorem 17.

Exercise 39. A density matrix ρ has rank 1 if and only if it corresponds to a pure state.

We collect some linear algebra notation, as we will now start relying on matrices (or linear operators) acting on vector spaces more heavily.[5]

Notation 40. Let $\mathcal{X}, \mathcal{Y}$ be two vector spaces. We define the following sets, relative to $\mathcal{X}, \mathcal{Y}$.

1. Linear operators.
   a) $\mathrm{L}(\mathcal{X}, \mathcal{Y})$: The set of liner operators from $\mathcal{X}$ to $\mathcal{Y}$.
   b) $\mathrm{L}(\mathcal{X})$: The set of linear operators from $\mathcal{X}$ to $\mathcal{X}$.
2. Unitary operators. (TODO)
3. Hermitian operators. $\mathrm{Herm}(\mathcal{X}) := \{M \in \mathrm{L}(\mathcal{X}) : M^\dagger = M\}$
4. Positive semidefinite operators. $\mathrm{Pos}(\mathcal{X}) := \{M \in \mathrm{Herm}(\mathcal{X}) : M \geq 0\}$ where $M \geq 0$ means that all its eigenvalues are non-negative.

2.3.2.2 Reduced states and Purifications

Consider a system register S and an environment register E. Suppose the state of the two registers initially is given by ρ_{SE}. How should one describe the state of register S by itself?

Stated in a more operational sense, suppose Aman prepares initialises registers SE into some initial state ρ_{SE} and gives register S to Basanti. How should Basanti describe the state of register S?

It turns out that the correct operation that captures this situation, is the partial trace operation defined as follows.

[5] We use the terms linear operator and matrices interchangeably as they are equivalent in finite dimensions.

Definition 41 (Partial Trace)**.** Let $M_{\mathsf{SE}} \in \mathrm{L}(\mathcal{S} \otimes \mathcal{E})$ be a matrix. Then,

$$\mathsf{tr}_{\mathsf{E}}(M_{\mathsf{SE}}) := \sum_{i \in I} \langle e_i| M_{\mathsf{SE}} |e_i\rangle$$

where $\{|e_i\rangle\}_{i \in I}$ is any orthonormal basis for $\mathcal{E}$. Equivalently,

$$\mathsf{tr}_{\mathsf{E}} \left(|s\rangle \langle s'|_{\mathsf{S}} \otimes |e\rangle \langle e'|_{\mathsf{E}}\right) = |s\rangle \langle s'|_{\mathsf{S}} \, \mathsf{tr}(|e\rangle \langle e'|)$$

for any set of vectors $|s\rangle, |s'\rangle \in \mathcal{S}$ and $|e\rangle, |e'\rangle \in \mathcal{E}$.

Notation 42 (Reduced Density matrix). When $\rho_{\mathsf{SE}} \in \mathrm{Pos}(\mathcal{S} \otimes \mathcal{E})$ is a density matrix, $\rho_{\mathsf{S}} := \mathsf{tr}_{\mathsf{E}}(\rho_{\mathsf{SE}})$ is termed the reduced density matrix.

Why is partial trace the correct operation here? It is not hard to see that we want the description ρ_{S} to be such that any observable acting non-trivially on register S should produce the same statistic, whether ρ_{S} is used or ρ_{SE} is. Partial trace satisfies this requirement.[6]

Exercise 43. Let AB be registers with $\mathcal{A}$ and $\mathcal{B}$ both $\mathbb{C}^2$. Write down the density matrix ρ_{AB} and the reduced density matrix ρ_{B} corresponding to the following pure states.

1. $|\psi\rangle_{\mathsf{A}} \otimes |\phi\rangle_{\mathsf{B}}$
2. $\sqrt{p_0}\, |0\rangle_{\mathsf{A}} \otimes |\psi_0\rangle_{\mathsf{B}} + \sqrt{p_1}\, |1\rangle \otimes |\psi_1\rangle$ where $p_0 \leq 1$ and $p_1 = 1 - p_0$.
3. $(|00\rangle + |11\rangle)_{\mathsf{AB}} / \sqrt{2}$

From Exercise 43, it is clear that pure states can give rise to mixed states when a part of the system is traced out. Can one do the opposite? Given a mixed state, can one always 'purify' it, i.e. given ρ_{A}, can one find $|\psi\rangle_{\mathsf{AB}}$ such that $\mathsf{tr}_{\mathsf{B}}(|\psi\rangle \langle\psi|_{\mathsf{AB}}) = \rho_{\mathsf{A}}$?

Exercise 44 (Purification)**.** Let $\rho_{\mathsf{A}} = \sum_{i \in I} \lambda_i \, |u_i\rangle \langle u_i|_{\mathsf{A}}$ be the spectral decomposition of a density matrix $\rho_{\mathsf{A}} \in \mathcal{A}$. Then, verify that $|\psi\rangle_{\mathsf{AB}} = \sum_{i \in I} \sqrt{\lambda_i}\, |u_i\rangle_{\mathsf{A}} \, |i\rangle_{\mathsf{B}}$ satisfies $\mathsf{tr}_{\mathsf{B}}(|\psi\rangle \langle\psi|_{\mathsf{AB}}) = \rho_{\mathsf{A}}$.

[6] For details, see page Box 2.6 of Nielsen and Chuang (10th edition).