PHYS 402 (Applications of Quantum Mechanics) Notes

Rio Weil

This document was typeset on October 31, 2024

Introduction:

This set of notes is based on UBC's PHYS 402 (Applications of Quantum Mechanics) course, taught by Dr. Robert Raussendorf. The course covers portions of chapters 1-5 of Sakurai's "Modern Quantum Mechanics" (though at the time of writing, only the first 2 chapters), as well as some additional topics on quantum entanglement. If any errors are found in the notes, feel free to email me at ryoheiweil@phas.ubc.ca.

Contents

1 Fundamental Concepts

1.1 The Beginnings of Quantum Mechanics

Before we dive headfirst into the formalism of quantum mechanics, let us first review the first steps of the field as taken in the early 1900s.

Our first founder is Max Planck; the problem at hand was the problem of the blackbody radiation spectrum. The two pre-existing laws (derived from thermodynamics arguments alone) predicting the BBR intensity as a function of wavelength/frequency were flawed. The first was Wien's law (1896):

$$I_{\text{Wien}}(\lambda, T) \sim \frac{1}{\lambda^5} \exp(-\frac{1}{\lambda T}) \tag{1.1}$$

which agreed with low wavelength/high frequency data well but failed to accurately describe high wavelength/low frequency emission. The second was Rayleigh-Jeans' law (1900):

$$I_{\text{RJ}}(\lambda, T) \sim \frac{T}{\lambda^4} \tag{1.2}$$

which agreed with high wavelength/low frequency data well but failed to accurately describe low wavelength/high frequency emission[1]. In fact, the intensity as predicted by Rayleigh-Jeans' diverges at low λ, leading to the (obviously) erroneous conclusion that the total energy emitted by a black body is infinite; the so-called "ultraviolet catastrophe".

In order to solve this problem, in 1900 Planck proposed a quantum hypothesis; that light carries energy in individual packets, or quanta. In particular, for light of frequency f, each quanta carries energy:

$$E = hf. \tag{1.3}$$

Combining this quantum hypothesis with the Boltzmann supression of high-energy states (from thermodynamics), Planck's law was then derived to be:

$$I_{\text{Planck}}(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp(\frac{hc}{\lambda k_B T}) - 1} \tag{1.4}$$

which agrees with the BBR spectrum data across all frequencies[2]. It should also be noted that the integral over all f of the above radiance law yields is finite, resolving the ultraviolet catastrophe.
In the above discussion, we have introduced Planck's constant. It has numerical value[3]:

$$h = 6.626070040 \times 10^{-34} \text{J s} \tag{1.5}$$

h is quantified as "small". What exactly does small mean in this context? For comparison, 1eV is the kinetic energy of an electron acquired in a voltage drop of a Volt, 0.035eV is the average kinetic energy of an atom at room temperature (from $E_k = \frac{3}{2} k_B T$) and 2.4eV is the energy of a single photon from the middle of the visible spectrum (600THz). The energy of a single photon, which depends on h, is in other words "typical" of microscopic phenomena.

Planck's quantum hypothesis would be confirmed in Einstein's (Nobel-prize winning) 1905 explanation of the photoelectric effect (which you likely covered in detail in a previous course in modern physics); namely that quanta of light transfer energy $E = hf$ to electrons in the metal, kicking them out[4].

[1] It should be noted however that a full-derivation of the Rayleigh-Jeans law did not occur until 1905, at which point Planck had already established the more correct explanation.

[2] Further, we can observe that Planck's law agrees with Wien's law in the high-frequency limit, and with Rayleigh-Jeans' law in the low-frequency limit.

[3] which is the set/absolute (rather than measured) value of the Planck constant as per the 2018 redefinition of SI units.

[4] Provided of course that $hf > \Phi$ where Φ is the "work function" of the metal.

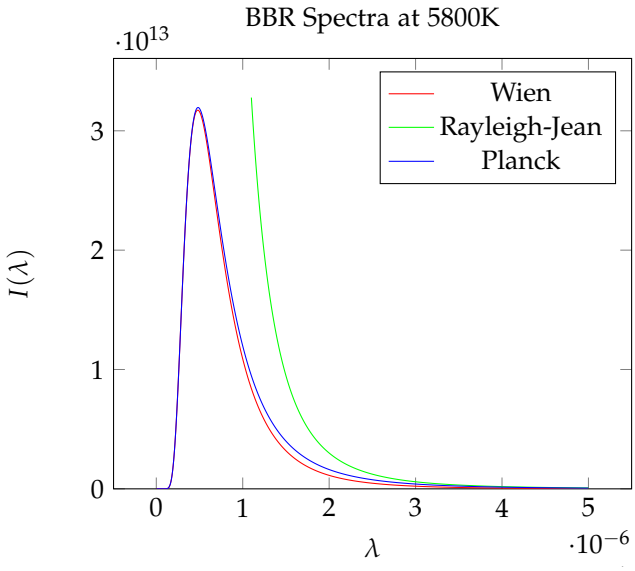

Figure 1.1: Plots of the black body emission spectra at $T = 5800$K (the approximate temperature of the surface of the sun) as predicted by Wien's Law, Rayleigh-Jean's Law, and Planck's Law. Planck's Law was found to agree with experimental observations for all wavelengths. Wien's Law agrees with observations well in the short wavelength limit but fails for long wavelengths. Rayleigh-Jean's Law agrees with observations in the long wavelength limit but fails at short wavelengths, and in fact the predicted emitted energy diverges.

Our second founder of interest is DeBroglie. In 1924, he postulated that matter could behave like a wave, positing the DeBroglie wavelength relation:

$$p = \frac{h}{\lambda}. \tag{1.6}$$

The so-called "wave-particle" duality would be confirmed in 1927 by the Davisson-Germer experiment, which saw peaks of electron intensity at distinct angles, showing that electrons scatter in the same nature as photons.

Our third founder of interest is Schrödinger, who postulated the Schrödinger equation (expressed below in the position basis) in 1926:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \left[\frac{-\hbar^2}{2m} \nabla^2 + V(\mathbf{r}, t) \right] \psi(\mathbf{r}, t). \tag{1.7}$$

It should be noted that this is one of the two core formulas of non-relativistic quantum mechanics, and is the quantum-mechanical equivalent of Newton's laws. It however does not cover the effects of special relativity (for which we defer the reader to a future course on quantum field theory) or quantum measurement (which we shall address now).

An illuminating demonstration of quantum measurement takes the form of the Stern-Gerlach experiment (first carried out in 1921/1922; see this article for more historical background). In this experiment, silver atoms are heated and escape from an oven with uniform velocity. The beam of atoms then pass through an inhomogenous magnetic field (generated by an asymmetric pair of magnetic pole pieces) where they are deflected, before hitting a screen where their position is recorded.

Why are silver atoms used for this experiment? Moreover, what exactly is being measured? For this, we consider a simplified model of the atom (which will suffice for the purposes of explaining this experiment).

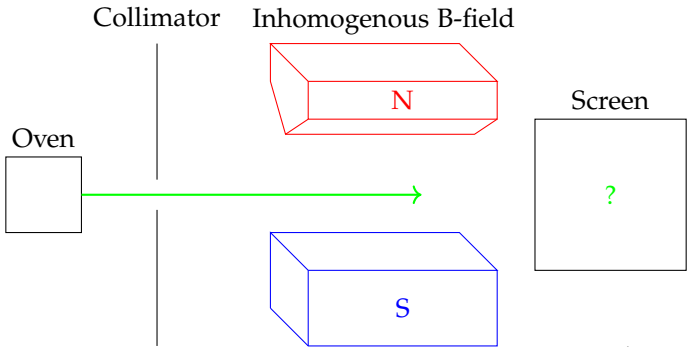

Figure 1.2: Illustration of the Stern-Gerlach experiment. Silver (Ag) atoms are heated in an oven and escape, and pass through a collimator to form a narrow beam. They then pass through an inhomogenous magnetic field which deflects the atoms. The position of the atoms is then recorded when they hit the screen.

Silver atoms consist of 47 electrons in the shell, and 47 protons and 61 neutrons in the nucleus. A first guess of the mechanism of the atoms being deflected by the magnetic field may be a Lorentz force effect; however this is not the case as the atoms are electrically neutral. Instead, the silver atom has a single unpaired electron which has an intrinsic angular momentum, known as spin. In particular, the electron is spin-1/2[5]. This provides the silver atom with a net magnetic moment $\boldsymbol{\mu}$ proportional to the electron spin[6] $\mathbf{S}$:

$$\boldsymbol{\mu} \propto \mathbf{S}. \tag{1.8}$$

We then recall from electromagnetism that a magnetic dipole $\boldsymbol{\mu}$ in a magnetic field $\mathbf{B}$ has interaction energy:

$$E = -\boldsymbol{\mu} \cdot \mathbf{B}. \tag{1.9}$$

We can then find the force that the dipole feels by taking the (negative) gradient of the energy:

$$\mathbf{F} = -\boldsymbol{\nabla}(-\boldsymbol{\mu} \cdot \mathbf{B}) = \begin{pmatrix} \dfrac{\partial}{\partial x}(\boldsymbol{\mu} \cdot \mathbf{B}) \\ \dfrac{\partial}{\partial y}(\boldsymbol{\mu} \cdot \mathbf{B}) \\ \dfrac{\partial}{\partial z}(\boldsymbol{\mu} \cdot \mathbf{B}) \end{pmatrix}. \tag{1.10}$$

Ignoring the magnetic fields that are not in the z-direction, we find the force on the silver atoms in the z-direction to be:

$$F_z = \mu_z \frac{\partial B_z}{\partial z}. \tag{1.11}$$

So in the inhomogenous field produced by the asymmetric magnets, the silver atoms should feel an up/downwards force depending on the direction of $\mathbf{S}$ (which determines μ_z).

Classically, the magnetic moment ¯ can point in any direction, and therefore μ_z ranges continuously from $+|\boldsymbol{\mu}|$ to $-|\boldsymbol{\mu}|$. Hence, the signature we would expect on the Stern-Gerlach experiment screen (wherein the vertical position of the atoms on the screen corresponds to a measurement of the z-component of the

[5]We will return to a more detailed discussion of angular momentum and spin at a later portion of the course

[6]The astute reader may question why the spin of the unpaired proton in the nucleus has no contribution to the net magnetic moment. This is due to the fact that the proportionality factor between the spin and magnetic moment has a factor of inverse mass. Since the proton is 1836 times heavier than the electron, the proton's magnetic moment contribution is negligeble compared to the electron's.

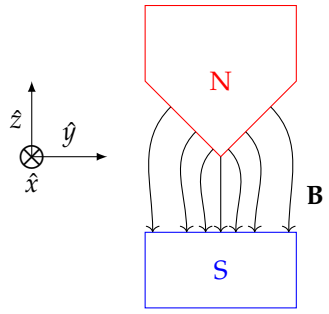

Figure 1.3: The inhomogenous magnetic field used in the Stern-Gerlach experiment, which deflects the silver atoms due to their magnetic dipole moment proportional to electron spin.

magnetic moment) would be a continuous band, as seen in the left of Fig. 1.4 below. However, this is *not* what is observed; instead the experimental result was two discrete dots with nothing in between, as seen in the right of the figure.

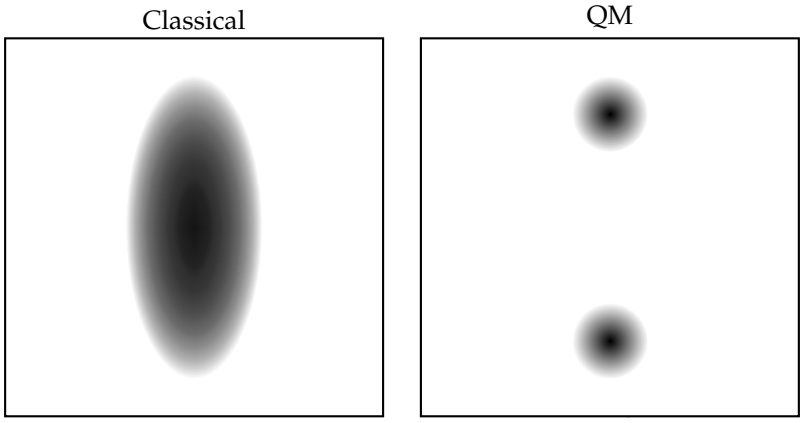

Figure 1.4: Classical prediction (left) and quantum mechanical prediction (right) for the Stern-Gerlach experiment. The screen on the right was observed in experiment.

How do we interpret this result? We can associate the top dot with spins fully polarized upwards ($\uparrow$) and the bottom dot with spins fully polarized downwards ($\downarrow$). But why is there no signature for sideways pointing spins? We first will answer how a general spin (1/2) state can be represented. If $|\uparrow\rangle$ represents the spin-up state and $|\downarrow\rangle$ represents the spin-down state, then a general spin (and hence sideways spins) can be represented as complex superpositions of these two states, i.e.

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle \tag{1.12}$$

where $\alpha, \beta \in \mathbb{C}$. What happens in a measurement is then that one element of this general superposition is picked with some probability; indeed, quantum measurement is a probabilistic process. Specifically, we find according to the Born rule that the probability that we measure the spin to be up is $p(\uparrow) = |\alpha|^2$ and the probability that we measure the spin to be down is $p(\downarrow) = |\beta|^2$. Since we require that we measure either spin-up or spin-down, we obtain the normalization condition:

$$p(\uparrow) + p(\downarrow) = |\alpha|^2 + |\beta|^2 = 1. \tag{1.13}$$

The spin state after the measurement is then $|\uparrow\rangle$ or $|\downarrow\rangle$ respectively, according to the Dirac projection

postulate. We will return to these two postulates of quantum mechanics and discuss them in full generality shortly.

However, we will however make a second comment about measurement before concluding this section. Namely, we consider the case where we perform a repeated measurement of the z-component of the spin. As discussed above, the initial general spin state is given by $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$. We then measure the z-component of spin and the post-measurement spin state is $|\uparrow\rangle$ or $|\downarrow\rangle$ with probability $|\alpha|^2$ and $|\beta|^2$ respectively. What happens if we measure the z-component of spin again? We might think that again, we have probability $|\alpha|^2$ of measuring spin-up and probability $|\beta|^2$ of measuring spin-down. But this is *not* the case. If we measured spin-up in the first measurement, we will measure spin-up in the second measurement with probability one. Similarly, if we measured spin-down in the first measurement, we will measure spin-down in the second measurement with probability one. Evidently, the first measurement has done something to the spin such that the measurement probabilities for the second measurement have been affected (they are not the same as the first). This tells us that quantum measurement is a active process that influences the state of the system we measure. Specifically, it is an irreversible process; there is no notion of "undo"-ing the measurement to recover the initial (pre-measurement) state.

1.2 Kets, Bras, and Hilbert Space

Our goal of the initial stages of this course will be to understand the following table:

Quantum states	$\|\psi\rangle \in \mathcal{H}$
Evolution	$i\hbar \frac{\partial}{\partial t}\|\psi\rangle = H\|\psi\rangle$
Measurement	$\|\psi\rangle \mapsto \frac{\Pi_j\|\psi\rangle}{\sqrt{\langle\psi\|\Pi_j\|\psi\rangle}} \quad p(j) = \langle\psi\|\Pi_j\|\psi\rangle$

Table 1: Axioms of quantum mechanics, concerning states, evolution, and measurement.

We will discuss the axioms for quantum states and quantum measurement in this chapter, and the axiom for quantum evolution (which readers may recognize as the Schrödinger equation in basis independent form) in the next. It is worth noting that these are the *fundamental postulates* of quantum mechanics; like Newton's laws of motion in classical mechanics, they cannot be derived. We are only able to interpret them, check if they are consistent, and work out the implications.

Let's start the axiom for quantum states; after all it will helpful to know what the objects of our interest are, before we start to work with them!

Axiom: Quantum states

Quantum states $|\psi\rangle$ are vectors (also called "kets") in a complex Hilbert space $\mathcal{H}$.

The above axiom is only meaningful if we know what a Hilbert space is; its definition is below:

Definition: (Complex) Hilbert spaces

$\mathcal{H}$ is a (complex) Hilbert space if:

(i) $\mathcal{H}$ is a vector space over $\mathbb{C}$

(ii) $\mathcal{H}$ has an inner product

(iii) $\mathcal{H}$ is complete (with respect to the metric induced by the norm induced by the inner product)[7]- For the purposes of this course, this last point can be ignored.

Note that the vector space axioms for closure imply that $\forall|\psi\rangle, |\varphi\rangle \in \mathcal{H}$ (where $\forall$ means "for all") and $\forall c \in \mathbb{C}$, then $|\psi\rangle + |\varphi\rangle \in \mathcal{H}$ and $c|\psi\rangle \in \mathcal{H}$. This tells us that the superposition of quantum states is well defined!

An example which we will return to time and time again (and have already encountered once) is the Hilbert space for a spin-1/2 system. In this case, $\mathcal{H} = \mathbb{C}^2$. A question that may be brooding in the reader's mind may be "why do we have to use complex numbers?"; one may indeed wonder if real Hilbert spaces may suffice to do quantum mechanics. The response is negative; we indeed need complex numbers! As an illustrative example, consider again the general spin-1/2 state $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$. Suppose we want a state that has equal probability to be measured spin-up or spin-down under a measurement of the z-component of spin. Since $p(\uparrow) = |\alpha|^2$ and $p(\downarrow) = |\beta|^2$, in order to have equal probability we must have $|\alpha| = 1/\sqrt{2}$ and $|\beta| = 1/\sqrt{2}$. A spin pointing in the $+x$ or $-x$ directions indeed has equal weight of up and down. Up to an overall (irrelevant) minus sign, without using complex numbers there are two ways to superimpose $|\uparrow\rangle$ and $|\downarrow\rangle$, from which we can define states corresponding to spins fully polarized in $\pm x$:

$$|x, \pm\rangle = |\rightarrow / \leftarrow\rangle = \frac{|\uparrow\rangle \pm |\downarrow\rangle}{\sqrt{2}}. \tag{1.14}$$

However, the $\pm\hat{x}$ and $\pm\hat{y}$ vectors lie in the same $z = 0$ plane, and by symmetry we should require that the $|y, \pm\rangle$ would also have equal weights of $|\uparrow\rangle$ and $|\downarrow\rangle$. But if we limit ourselves to real numbers only, we have already exhausted all possible equal-weight combinations of $|\uparrow\rangle$ and $|\downarrow\rangle$ in Eq. (1.14). We therefore require complex numbers to represent all possible states (and indeed, we find that $|y, \pm\rangle = \frac{|\uparrow\rangle \pm i|\downarrow\rangle}{2}$).

Having motivated the "complex" in the complex vector space part of the definition of Hilbert spaces, let us now motivate the inner product. We want some way to compare quantum states to one another. Our geometric intuition tells us that the states $|\uparrow\rangle$ and $|\nearrow\rangle$ are "close" to each other, while $|\uparrow\rangle$ and $|\downarrow\rangle$ are very "different". In order to make this intuition rigorous, we define the inner product, and as a prerequisite we define the dual correspondence.

Definition: Dual correspondence

To each vector space $\mathcal{H}$, there exists a dual vector space $\mathcal{H}^*$. There is a one-to-one correspondence[8]between the kets $|\psi\rangle \in \mathcal{H}$ and the bras[9] $\langle\psi| \in \mathcal{H}^*$. We call this the *dual correspondence*, and write it as follows:

$$|\psi\rangle \overset{DC}{\longleftrightarrow} \langle\psi|. \tag{1.15}$$

It has the following properties:

(i) $|\psi\rangle + |\varphi\rangle \overset{DC}{\longleftrightarrow} \langle\psi| + \langle\varphi|$

(ii) $c|\psi\rangle \overset{DC}{\longleftrightarrow} c^*\langle\psi|$

where the $*$ denotes complex conjugation.

Having established the dual correspondence, we may now define the inner product:

[7]This is a technical qualification for the mathematicians in the crowd. An intuitive explanation for the curious; the inner product on a Hilbert spaces creates a notion of distance on the space. There are sequences (of vectors) that get closer together over time; completeness tells us that any such sequences (known as Cauchy sequences) must converge to a limit.

[9]Formally, this follows from the Riesz Representation Theorem. But for the purposes of this course, we take this one-to-one correspondence as a postulate. Curious readers can find discussions/proofs of the theorem in any text on functional analysis, or mathematical quantum theory.

[9]Given such names because $\langle|\rangle$ is a bracket - bra-ket. Physicists remain unmatched in their sense of humour.

Definition: Inner product

We define the inner product between $|\psi\rangle \in \mathcal{H}$ and $|\varphi\rangle \in \mathcal{H}$ as:

$$\langle\varphi|\psi\rangle \in \mathbb{C} \tag{1.16}$$

with the properties:

(i) $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$

(ii) $\langle\psi|\psi\rangle \geq 0, \forall|\psi\rangle \in \mathcal{H}$

(iii) $\langle\psi|\psi\rangle = 0 \implies |\psi\rangle = \mathbf{0}$.

Note that $\mathbf{0}$ in the above definition is the null ket (also known as the zero vector, which must be an element of the Hilbert space), where $\mathbf{0} = 0|\psi\rangle, \forall|\psi\rangle$. It is an unphysical state. We normally work with normalized states, i.e. states $|\psi\rangle$ that satisfy $\langle\psi|\psi\rangle = 1$. The null ket has inner product zero and cannot be normalized.

As a first use of the inner product, let us return to our initial motivation for obtaining the "likeness" of states. For normalized states, it follows (and we will later prove) that:

$$0 \leq |\langle\varphi|\psi\rangle| \leq 1. \tag{1.17}$$

We therefore can use the inner product as a method to evaluate the likeness of states. $\langle\varphi|\psi\rangle = 0$ means that $|\psi\rangle$ and $|\varphi\rangle$ are maximally different, and $|\langle\varphi|\psi\rangle| = 1$ corresponds to $|\psi\rangle$ and $|\varphi\rangle$ being the same.

Next, we move onto a discussion of bases of Hilbert spaces. Since Hilbert spaces are vector spaces, they admit a basis. Let us recall what a basis is.

Definition: Basis

A *basis* $\mathcal{B}$ of $\mathcal{H}$ is a set of states $\mathcal{B} = \left\{|b_j\rangle\right\}_j$ such that every state $|\psi\rangle \in \mathcal{H}$ can be written in the form:

$$|\psi\rangle = \sum_j \psi_j |b_j\rangle \tag{1.18}$$

with $\psi_j \in \mathbb{C} \forall j$, and the expansion on the RHS is unique.
$|\mathcal{B}|$ is the dimension of $\mathcal{H}$[10].

Of particular interest to us will be orthonormal bases, or ONBs.

[10] In the case where the Hilbert space is infinite-dimensional, there are additional complications (and in fact, two different kinds of bases!) In general the work we do in this course works perfectly well in the finite-dimensional case, and in the infinite dimensional case we will have to wave our hands a bit in order to avoid getting into the weeds of functional analysis; this is a physics course after all.

Definition: Orthonormal bases

A basis $\mathcal{B} = \left\{|b_j\rangle\right\}_j$ is *orthonormal* if

$$\langle b_i|b_j\rangle = \delta_{ij}. \tag{1.19}$$

where δ_{ij} is the Kronecker delta, defined as:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \tag{1.20}$$

One can obtain the expansion coefficients ψ_j with respect to ONBs. Writing $|\psi\rangle = \sum_j \psi_j|b_j\rangle$ as in Eq. (1.18) and taking the inner product of $|\psi\rangle$ with a vector $|b_i\rangle$ from the ONB, we have:

$$\langle b_i|\psi\rangle = \sum_j \psi_j\langle b_i|b_j\rangle = \sum_j \psi_j\delta_{ij} = \psi_i. \tag{1.21}$$

We will now prove a useful trick involving ONBs.

Proposition: Resolution of the identity

For all ONBs $\left\{|b_j\rangle\right\}_j$, the following relation holds:

$$\sum_j |b_j\rangle\langle b_j| = \mathbb{I} \tag{1.22}$$

where $\mathbb{I}$ is the identity operator on the Hilbert space.

Proof. Recall that $|\psi\rangle = \sum_j \psi_j|b_j\rangle$ for any $|\psi\rangle \in \mathcal{H}$ and for any basis $\left\{|b_j\rangle\right\}_j$ of $\mathcal{H}$. Further, recall that $\psi_j = \langle b_j|\psi\rangle$ if the basis is orthonormal. Hence we have that:

$$|\psi\rangle = \sum_j \langle b_j|\psi\rangle|b_j\rangle = \sum_j |b_j\rangle\left(\langle b_j|\psi\rangle\right) = \sum_j \left(|b_j\rangle\langle b_j|\right)|\psi\rangle = \left(\sum_j |b_j\rangle\langle b_j|\right)|\psi\rangle. \tag{1.23}$$

Since the above relation holds for all $|\psi\rangle$, it follows then that $\sum_j |b_j\rangle\langle b_j|$ is the identity as claimed. □

At this point in the course, the reader may be wondering what happened to quantum *wavefunctions*[11]; the central objects of interest have instead been quantum states, without a wavefunction in sight. We now elucidate the connection between the two.

[11] Though this nomenclature of "wavefunction" is arguably a misnomer; the Schrödinger equation does not contain second order derivatives in time, as a wave equation would.

Definition: Wavefunctions

Consider (as before) expanding $|\psi\rangle$ in a ONB $\left\{|b_j\rangle\right\}_j$. We then have:

$$|\psi\rangle = \sum_j |b_j\rangle\langle b_j|\psi\rangle. \tag{1.24}$$

Therein, $\psi_j = \psi(j) = \langle b_j|\psi\rangle$ is the *wavefunction*; a wavefunction is just a quantum state expanded in a given ONB.

A particularly important (and familiar) example is using the position basis. Consider the resolution of the identity involving position:

$$\int \mathrm{d}x|x\rangle\langle x| = \mathbb{I}. \tag{1.25}$$

For any $|\psi\rangle$, we then have:

$$|\psi\rangle = \int \mathrm{d}x|x\rangle\langle x|\psi\rangle. \tag{1.26}$$

Where $\langle x|\psi\rangle = \psi(x)$ is the position wavefunction, which played a central role in your first course in quantum mechanics. However, the reader should now recognize that $|\psi\rangle$ is a more fundamental object than this wavefunction, as it not only contains the information for $\psi(x)$ but also for $\tilde{\psi}(p)$ (the momentum wavefunction) or any other wavefunction; any wavefunction is just the expansion coefficients of the state in a given basis.

We have now established quite a bit of machinery to discuss quantum states, but have not done anything with the states; let's change that by discussing measurements! In a quantum measurement, the alternatives $\left\{|m_j\rangle, j \in \text{ Outcomes }\right\}$ for the states after measurement will:

- Form a basis
- And are pairwise orthogonal, i.e. $\langle m_i|m_j\rangle = 0, \forall i \neq j$.

With the necessary mathematical formalism under our belt, we can now state a first version of the Dirac postulate and Born rule axioms.

Axiom: Quantum measurement (version 1)

Dirac postulate: Under quantum measurement, the measured quantum state $|\psi\rangle$ is probabilistically changed into one of a number of alternatives $\left\{|m_j\rangle\right\}_j$, with:

$$\langle m_i|m_j\rangle = \delta_{ij}. \tag{1.27}$$

Note that $\left\{|m_j\rangle\right\}_j$ forms an ONB.

Born rule: Given a quantum state $|\psi\rangle$, the probability for obtaining outcome j, corresponding to post-measurement state $|m_j\rangle$, is:

$$p_j = |\langle m_j|\psi\rangle|^2. \tag{1.28}$$

There are now two questions that may arise. The first is that the statement of the Dirac projection postulate and the Born rule do not match that found in Table 1. The second question concerns the post-measurement states $\left\{|m_j\rangle\right\}_j$; namely, how are they are determined? We will address the latter question first, and build up the formalism to state the measurement axioms in full generality. To this end, we move to a discussion of linear operators, which describe quantum mechanical observables and evolution.

1.3 Operators and Observables

Definition: Linear operators

A is an *operator* that acts on a Hilbert space $\mathcal{H}$ if $\forall|\psi\rangle \in \mathcal{H}$, $A|\psi\rangle \in \mathcal{H}$. A is *linear* if:

$$A(\alpha|\psi\rangle + \beta|\varphi\rangle) = \alpha(A|\psi\rangle) + \beta(A|\varphi\rangle) \tag{1.29}$$

$\forall|\psi\rangle, |\varphi\rangle \in \mathcal{H}$ and $\forall\alpha, \beta \in \mathbb{C}$.

A point of notation; we will use capital letters to denote operators in this course. Some sources also use hats to denote this (e.g. $\hat{A}$). Note that linear operators can be added and multiplied (more accurately, composed) to yield other linear operators. They are associative and distributive under these operations, i.e. for all linear operators A, B, C we have:

$$(A + B) + C = A + (B + C) \tag{1.30}$$

$$(AB)C = A(BC) \tag{1.31}$$

$$A(B + C) = AB + AC. \tag{1.32}$$

However, note that in general operators are *not* commutative, that is:

$$AB \neq BA. \tag{1.33}$$

Quantum mechanical observables are a specific type of linear operators; namely, they are Hermitian. In order to make sense of this, we introduce two more definitions.

Definition: Hermitian adjoint

The *Hermitian adjoint* $A^\dagger$ of a linear operator A is defined via the dual correspondence:

$$A|\psi\rangle \overset{DC}{\longleftrightarrow} \langle\psi|A^\dagger. \tag{1.34}$$

Definition: Hermitian operators and observables

An operator A is *Hermitian* if:

$$A^\dagger = A. \tag{1.35}$$

Quantum mechanical *observables* (such as position, momentum, spin, and energy) are Hermitian operators.

In general, acting on a state with an operator changes the state, and the new state is not necessarily proportional to the original state, i.e:

$$A|\psi\rangle \not\propto |\psi\rangle. \tag{1.36}$$

However, for special states known as eigenstates of an operator, this is indeed the case.

Definition: Eigenstates and eigenvalues

An *eigenstate* $|a\rangle$ of a linear operator A is a state that satisfies:

$$A|a\rangle = a|a\rangle. \tag{1.37}$$

Therein, $a \in \mathbb{C}$ is the *eigenvalue* corresponding to that eigenstate. The null vector $\mathbf{0}$ is excluded from being an eigenvector.

Having defined these objects, we state (but do not prove) an important theorem concerning observables:

Theorem: Diagonalization/Spectral theorem

A Hermitian operator A (i.e. any observable) can be diagonalized; that is, it is able to be written as:

$$A = \sum_i a_i |a_i\rangle\langle a_i| \tag{1.38}$$

Where the eigenvectors $\{|a_i\rangle\}$ of A forms an orthonormal basis[12](an *eigenbasis*) and a_i are the corresponding eigenvalues.

With some more machinery developed, we can more carefully state the Dirac postulate and the Born rule.

Axiom: Quantum measurement (version 2)

Dirac postulate:

1. Each possible outcome of the measurement of an observable A is an eigenvalue of A.
2. If the eigenvalue a is found in the measurement, then the post measurement state is an eigenvector $|a\rangle$ of A,
$$|\psi\rangle \mapsto |a\rangle \tag{1.39}$$
satisfying
$$A|a\rangle = a|a\rangle. \tag{1.40}$$

Born rule: Given an initial state of $|\psi\rangle$, the possibility for the measurement outcome a_i occuring in a measurement of an observable A (where a_i is an eigenvalue of A) is:

$$p_i = |\langle a_i|\psi\rangle|^2. \tag{1.41}$$

We reiterate that the eigenstate $|a\rangle$ is the possible post-measurement state, and the eigenvalue a is the corresponding measurement outcome. As a concrete example, the spin-z operator S_z (for spin-1/2) systems has eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$, with corresponding eigenvalues of $+\hbar/2$ and $-\hbar/2$ (i.e. $S_z|\uparrow\rangle = +\hbar/2|\uparrow\rangle$, and likewise for $|\downarrow\rangle$). We may write S_z in diagonal form as:

$$S_z = \frac{\hbar}{2}(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|). \tag{1.42}$$

A measurement of the z component of spin has two possible outcomes $\pm\hbar/2$, corresponding to post measurement states $|\uparrow\rangle/|\downarrow\rangle$. $\{|\uparrow\rangle, |\downarrow\rangle\}$ is an ONB, and we can expand a general state $|\psi\rangle$ in this basis as $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$. The Born rule then tells us that $p(\uparrow) = |\langle\uparrow|\psi\rangle|^2 = |\alpha|^2$ and similarly that $p(\downarrow) = |\beta|^2$.

There are two points of consistency that the restatement of the Dirac postulate invites. First, an experiment should only have real-valued outcomes; an measurement shouldn't return a complex number. Second, it is not a priori immediate that the eigenstates/post-measurement states are mutually orthogonal, as the first statement of the postulate requires. Fortunately, there is a theorem that covers both.

[12]If the eigenvalues of A are non-degenerate, then the set of eigenstates of A are automatically mutually orthogonal, as we prove in the next theorem. If A has degenerate eigenvalues, two eigenstates with the same eigenvalue are not necessarily guaranteed to be orthogonal, but we are still able to choose eigenstates (picking orthogonal vectors from the degenerate subspaces) to form an ONB out of the eigenstates; so we need not worry too much.

Theorem

The eigenvalues of a Hermitian operator A are real, and the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof. Consider two eigenstates $|a\rangle, |a'\rangle$ of A (with corresponding eigenvalues a, a'). First, we make the observation that $\langle a|A = \langle a|a^*$. This follows as:

$$\langle a|a^* \overset{DC}{\longleftrightarrow} a|a\rangle = A|a\rangle = A^\dagger|a\rangle \overset{DC}{\longleftrightarrow} \langle a|A \tag{1.43}$$

where the first equality invokes that $|a\rangle$ is an eigenvector of A, and the second equality invokes the Hermicity of A. Since the dual correspondence is unique, we can compare the first and the last expressions to conclude that $\langle a|A = \langle a|a^*$. Next, we consider the number $\langle a|A|a'\rangle$. We may write this as:

$$\langle a|A|a'\rangle = \langle a|(A|a'\rangle) = \langle a|a'|a\rangle = a'\langle a|a'\rangle \tag{1.44}$$

but by associativity, we can just as well write this as:

$$\langle a|A|a'\rangle = (\langle a|A)|a'\rangle = \langle a|a^*|a'\rangle = a^*\langle a|a'. \tag{1.45}$$

Therefore we find that $a^*\langle a|a' = a'\langle a|a'\rangle$. From this we obtain that:

$$(a^* - a')\langle a|a'\rangle = 0. \tag{1.46}$$

There are now two cases to consider.

(I) If $|a\rangle = |a'\rangle$, then $\langle a|a'\rangle = \langle a|a\rangle > 0$ (as $|a\rangle$ is an eigenvector, it cannot be a null vector). Therefore for Eq. (1.46) to be satisfied it must follow that $a^* = a$, i.e. a is real.

(II) If instead $a \neq a'$, then $a^* \neq a'$ (as $a = a^*$), so for Eq. (1.46) to be satisfied it must follow that $\langle a|a'\rangle = 0$.

□

We have now shown that the Dirac projection postulate is consistent with what we should expect out of measurements. However, we now clarify what is often a point of confusion, that being the difference between individual and averaged measurement outcomes. The dirac postulate states that given a state $|\psi\rangle$, a possible outcome of the measurement of an observable A are eigenvalues a of A. This speaks to possible outcomes in *individual measurements*. The *expectation value* is conceptualized quite differently.

Definition: Expectation values

The *expectation value* $\langle A\rangle_\psi$ is the average outcome in the measurement of A on $|\psi\rangle$, i.e.:

$$\langle A\rangle_\psi := \sum_i p_i a_i \tag{1.47}$$

where p_i is the probability of measuring outcome a_i. We are able to write Eq. (1.47) equivalently as:

$$\langle A\rangle_\psi = \sum_i |\langle a_i|\psi\rangle|^2 a_i = \sum_i \langle a_i|\psi\rangle^* \langle a_i|\psi\rangle a_i = \sum_i \langle\psi|a_i\rangle a_i \langle a_i|\psi\rangle = \langle\psi|\left(\sum_i a_i|a_i\rangle\langle a_i|\right)|\psi\rangle = \langle\psi|A|\psi\rangle. \tag{1.48}$$

where in the first equality we use the Born rule, and in the last equality we consider A in diagonalized form.

The formalism of quantum mechanics is therefore well-suited to predict both individual and averaged measurement outcomes; but take care not to confuse them! In general $\langle A \rangle_\psi$ is *not* a possible (individual) measurement outcome.

Up until now, we have considered both kets $|\psi\rangle$ and operators A as abstract objects; however sometimes it is helpful to cast these into a more concrete form in order to do computations. We thus introduce the idea of matrix/vector representations.

Consider an ONB $\mathcal{B} = \{|b_i\rangle\}_i$, and consider an expression of the form $|\varphi\rangle = A|\psi\rangle$. Inserting the resolution of the identity once on the left hand side and twice on the right hand side, we have:

$$\left(\sum_i |b_i\rangle\langle b_i|\right)|\varphi\rangle = \left(\sum_i |b_i\rangle\langle b_i|\right) A \left(\sum_j |b_j\rangle\langle b_j|\right)|\psi\rangle \tag{1.49}$$

Now redrawing some brakets we obtain:

$$\sum_i |b_i\rangle(\langle b_i|\varphi\rangle) = \sum_i \sum_j |b_i\rangle(\langle b_i|A|b_j\rangle)(\langle b_j|\psi\rangle). \tag{1.50}$$

Now, we can consider $\langle b_i|\varphi\rangle = [\varphi]_i$ and $\langle b_j|\psi\rangle = [\psi]_j$ (where we use $[\cdot]$ to denote 'representation of') as elements of column vectors $[\varphi]$ and $[\psi]$, and $\langle b_i|A|b_j\rangle = [A]_{ij}$ as elements of a matrix $[A]$. In other words, expanded out in the $\{|b_i\rangle\}_i$ basis, we can realize $[\varphi]$ is the column vector obtained by multiplying the column vector $[\psi]$ by the matrix $[A]$.

Similarly, we can write for any bra $\langle\tau|$ that:

$$\langle\tau| = \langle\tau|\sum_i |b_i\rangle\langle b_i| = \sum_i \langle\tau|b_i\rangle\langle b_i| = \sum_i [\tau]_i \langle\tau|. \tag{1.51}$$

This yields us the following way of thinking about bras/kets/operators:

Abstract Objects	Representation in ONB
Ket $\|\varphi\rangle$	Column vector $[\varphi]$
Operator A	Matrix $[A]$
Bra $\langle\tau\|$	Row vector $[\tau]$

Table 2: Abstract objects and their representations when expanded out in a basis.

However, do take note that a ket is *not* equal to a column vector. The column vector is a representation of a ket, much in the same way that $\binom{1}{0}$ is not in itself a vector but instead a concrete (standard) representation of the abstract vector e_1.

Let us supplement this discussion of matrix representations by returning to our favourite example of spin-1/2. Consider the ONB $\mathcal{B} = \{|\uparrow\rangle, |\downarrow\rangle\}$ of $\mathcal{H} = \mathbb{C}^2$. Making the identification that $|\uparrow\rangle \cong \binom{1}{0}$ and $|\downarrow\rangle \cong \binom{0}{1}$, for any $|\psi\rangle\mathcal{H}$ we then have that:

$$|\psi\rangle \cong \begin{pmatrix} \langle\uparrow|\psi\rangle \\ \langle\downarrow|\psi\rangle \end{pmatrix} = [\psi] \tag{1.52}$$

and for any linear operator A acting on kets in $\mathcal{H}$ we have:

$$A \cong \begin{pmatrix} \langle\uparrow|A|\uparrow\rangle & \langle\uparrow|A|\downarrow\rangle \\ \langle\downarrow|A|\uparrow\rangle & \langle\downarrow|A|\downarrow\rangle \end{pmatrix} = [A]. \tag{1.53}$$

Taking spin-z operator S_z as a concrete example, we can represent it in this basis as:

$$S_z = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\right) \cong \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.54}$$

As another example, consider the spin-x operator S_x. It has eigenstates:

$$|\rightarrow / \leftarrow\rangle := \frac{|\uparrow\rangle \pm |\downarrow\rangle}{\sqrt{2}} \tag{1.55}$$

we can represent it as:

$$S_x = \frac{\hbar}{2}\left(|\rightarrow\rangle\langle\rightarrow| - |\leftarrow\rangle\langle\leftarrow|\right) \cong \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{1.56}$$

From the work we have done so far, it may have become clear that it is generally easiest to work in the eigenbasis of whatever observable is being considered. For example, if considering the measurement of the observable S_z, it is natural to work with the basis $\mathcal{B} = \{|\uparrow\rangle, |\downarrow\rangle\}$. This fact will continue to be true when we later consider Schrödinger evolution of observables. However, there are times when there are multiple observables under consideration; in the measurement picture, this could be the sequential measurement of observable O_1 followed by the measurement of observable O_2[13]. In this case, we have two different bases of interest, namely the eigenbasis of O_1 and the eigenbasis of O_2. We may then find it useful to consider a transformation between these bases; let us work through that now.

Consider two ONBs defined by $\mathcal{B} = \{|b_i\rangle\}_i$ and $\mathcal{B}' = \{|a_j\rangle\}_j$, and some state $|\psi\rangle \in \mathcal{H}$. Inserting the resolution of identity $\mathbb{I} = \sum_j |a_j\rangle\langle a_j|$ between the wavefunction $\langle b_i|\psi\rangle = \psi_i$, we obtain:

$$\langle b_i|\psi\rangle = \langle b_i|\left(\sum_j |a_j\rangle\langle a_j|\right)|\psi\rangle = \sum_j \langle b_i|a_j\rangle\langle a_j|\psi\rangle \tag{1.57}$$

Now, $\langle b_i|\psi\rangle$ can be viewed as elements of a column vector $[\psi]$ (whose components are the expansion coefficients of $|\psi\rangle$ in the basis $\mathcal{B}$), $\langle a_i|\psi\rangle$ can be analogously viewed as the elements of a column vector $[\tilde{\psi}]$ (whose components are the expansion coefficients of $|\psi\rangle$ in the basis $\mathcal{B}'$), and $\langle b_i|a_j\rangle = T_{ij}$ can be viewed as the matrix elements of a (matrix) representation of the transformation operator T between the two bases. We can therefore write Eq. (1.57) as a matrix/vector equation:

$$[\psi] = [T][\tilde{\psi}]. \tag{1.58}$$

The operator T actually has a special property; namely, it is *unitary*.

Definition: Unitary operators

A linear operator U is *unitary* if:

$$UU^\dagger = U^\dagger U = \mathbb{I}. \tag{1.59}$$

Proposition

The basis transformation operator between bases $\mathcal{B} = \{|b_i\rangle\}_i$ and $\mathcal{B}' = \{|a_j\rangle\}_j$ defined as:

$$T = \sum_n |a_n\rangle\langle b_n| \tag{1.60}$$

which has matrix representation (in either $\mathcal{B}$ or $\mathcal{B}'$):

$$[T]_{ij} = \langle b_i|a_j\rangle \tag{1.61}$$

is unitary.

[13] An similar setting of interest with Schrödinger evolution could be the evolution of a quantum state under a Hamiltonian H_1 for some time $[t_0, t_1]$ followed by evolution by a different Hamiltonian H_2 for some time $[t_1, t_2]$.

Proof. First, it should be clear from the basis agnostic definition in Eq. (1.60) that T is in fact a basis transformation operator, as by inspection it satisfies $T|b_n\rangle = |a_n\rangle$ for $n = 1, 2, \ldots d$. We next verify that it indeed has the claimed matrix representation; in the $\mathcal{B}$ basis, we have:

$$[T]_{ij} = \langle b_i|T|b_j\rangle = \langle b_i| \left(\sum_n |a_n\rangle\langle b_n| \right) |b_j\rangle = \sum_n \langle b_i|a_n\rangle\langle b_n|b_j\rangle = \sum_n \langle b_i|a_n\rangle\delta_{nj} = \langle b_i|a_j\rangle \tag{1.62}$$

and in the $\mathcal{B}'$ basis we have:

$$[T]_{ij} = \langle a_i|T|a_j\rangle = \langle a_i| \left(\sum_n |a_n\rangle\langle b_n| \right) |a_j\rangle = \sum_n \langle a_i|a_n\rangle\langle b_n|a_j\rangle = \sum_n \delta_{in}\langle b_n|a_j\rangle = \langle b_i|a_j\rangle \tag{1.63}$$

so Eq. (1.61) indeed holds as claimed.

We move onto the unitarity proof. We claim that the equality below:

$$[T^\dagger]_{kl} = [T]^*_{lk} \tag{1.64}$$

holds *in general* for any linear operator T. Note that Eq. (1.64) reconciles the abstract notion of the Hermitian adjoint with the familiar operation of taking the conjugation transpose of a matrix (which you should have encountered in your linear algebra course). The equation follows from the definition of the Hermitian adjoint using the dual correspondence, and is left as an exercise for the reader.

Applying Eq. (1.64) to the definition of T, we find:

$$[T^\dagger]_{kl} = [T]^*_{lk} = \langle b_l|a_k\rangle^* = \langle a_k|b_l\rangle. \tag{1.65}$$

Next, we consider the matrix elements of the operator $T^\dagger T$:

$$[T^\dagger T]_{kj} = \sum_l [T^\dagger]_{kl}[T]_{lj} = \sum_l \langle a_k|b_l\rangle\langle b_l|a_j\rangle = \langle a_k| \left(\sum_l |b_l\rangle\langle b_l| \right) |a_j\rangle. \tag{1.66}$$

where in the first equality we use the definition of matrix multiplication, and in the second equality we use Eq. (1.65). Recognizing the resolution of the identity, the above equation then becomes:

$$[T^\dagger T]_{kj} = \langle a_k|\mathbb{I}|a_j\rangle = \langle a_k|a_j\rangle = \delta_{kj} \tag{1.67}$$

where in the last equality we use that $\left\{|a_j\rangle\right\}_j$ is an ONB. From $[T^\dagger T]_{kj} = \delta_{kj}$ we can conclude that $T^\dagger T = \mathbb{I}$, and $TT^\dagger = \mathbb{I}$ can be shown analogously. We conclude that T is unitary. □

Note that while this is our first encounter with unitarity, it will certainly not be our last; unitary operators will have a very large role to play when we consider Schrödinger evolution.

As an example, let us solve for T which transforms from the S_z eigenbasis $\mathcal{B} = \{|\uparrow\rangle, |\downarrow\rangle\}$ to the S_x eigenbasis $\mathcal{B} = \left\{|\rightarrow\rangle = \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}, |\leftarrow\rangle = \frac{|\uparrow\rangle-|\downarrow\rangle}{\sqrt{2}}\right\}$. We may write it down in operator form as:

$$T = |\rightarrow\rangle\langle\uparrow| + |\leftarrow\rangle\langle\downarrow|. \tag{1.68}$$

To get a clearer picture, we can write down its matrix representation. Computing the inner products between the eigenvectors of S_z and S_x, we obtain:

$$[T] = \begin{pmatrix} \langle\uparrow|\rightarrow\rangle & \langle\uparrow|\leftarrow\rangle \\ \langle\downarrow|\rightarrow\rangle & \langle\downarrow|\leftarrow\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{1.69}$$

The above matrix representation is the same whether we choose the identification $|\uparrow\rangle \cong \binom{1}{0}, |\downarrow\rangle \cong \binom{0}{1}$ or $|\rightarrow\rangle \cong \binom{1}{0}, |\leftarrow\rangle \cong \binom{0}{1}$ (but note: this is NOT true of operators in general. For example the matrix representations of S_z and S_x will look different depending on the identifications chosen). It can be easily verified via matrix multiplication that T is unitary, which is consistent with the general theorem we just proved. The T above is a ubiquitous operator in the field of quantum computation, and is called the *Hadamard operator*.

1.4 Projectors and Measurement

We now return to the setting of measurement to resolve an ambiguity that was left by version 2 of our measurement axioms. Namely, we ask "what if the observable in question has degenerate eigenvalues?" (Degenerate referring to the fact that two or more distinct eigenstates of an observable may have the same eigenvalues). To see why this is a problem with our current formulation, consider an operator A with two distinct eigenstates $|a\rangle, |\tilde{a}\rangle$ such that:

$$A|a\rangle = a|a\rangle, A|\tilde{a}\rangle = a|\tilde{a}\rangle. \tag{1.70}$$

In our earlier formulation of the Dirac postulate, we claimed that a measurement of an observable A with outcome/eigenvalue a would change the quantum state to be an eigenstate with the corresponding eigenvalue. Now that there are two distinct eigenstates with the same eigenvalue, which eigenstate is chosen?

As a concrete example, we consider the z-component of spin for a spin-1 particle. The S_z operator has three eigenstates, namely $|+\rangle, |0\rangle$ (this is *not* the null ket/zero vector!) and $|-\rangle$ which forms a basis, and can be written as:

$$S_z = \hbar(|+\rangle\langle+| - |-\rangle\langle-|). \tag{1.71}$$

$|+\rangle$ has eigenvalue $+\hbar$, $|0\rangle$ has eigenvalue zero, and $|-\rangle$ has eigenvalue $-\hbar$. Now, consider the observable S_z^2. From S_z above, we can deduce this to be:

$$S_z^2 = \hbar^2(|+\rangle\langle+| + |-\rangle\langle-|) \tag{1.72}$$

Now, we find that both $|+\rangle$ and $|-\rangle$ have eigenvalue $+\hbar^2$. So, if we measured S_z^2 for a state $|\psi\rangle$ and got outcome $\hbar^2$, our current formulation is ill-equipped to deduce what the post-measurement state would be. In order to refine our formulation, we require the notion of a projector.

Another example (which we cannot discuss in full detail yet) is that of "spooky action at a distance" (coined by the one and only Einstein). In this setting, we have a two-particle system which is in an entangled Bell state (to be defined in full mathematical detail later; for now, take entangled to mean "stronger than classically correlated", or "it is impossible to describe the individual subsystems/particles, and only the joint system can be described"). We then separate the two particles some great distance; say the Earth and moon. When we measure the particle on Earth (but do nothing to the particle on the moon), we find that upon knowing the Earth measurement outcome we *immediately* know what the post-measurement state of the particle on the moon is. This seems counterintutive, and indeed, superliminal; it seems as though the measurement on Earth affects the state on the moon faster than the speed of light! While this setup does not actually turn out to violate special relativity (and indeed, we will prove later in the course that no quantum communication protocol can be superliminal), it is certainly of interest as depicting non-classical phenomenology. It turns out that the measurement of just the particle on Earth conducted in this setting corresponds to a measurement of a degenerate (joint) observable; we therefore also require a reformulation of our measurement axioms involving projectors here. Without further ado, let us define what a projector is.

Figure 1.5: Visualization of the "spooky action at a distance" scenario. A two-particle state is prepared in an entangled Bell state. The particles are then spatially separated a great distance; here depicted as one particle on the Earth and one particle in the moon. When the particle on Earth (and *only* the particle on Earth) is measured, the entanglement structure of the state leads to the particle on the moon being projected into a post-measurement state that is immediately known to the experimenter on Earth based on their measurement outcome. The measurement conducted in this scenario of interest turns out to be one with degeneracy, and hence we require a reformulation of measurement in order to describe it.

Definition: Projectors

A linear operator Π is a *projector* if it satisfies:

$$\Pi^2 = \Pi^\dagger = \Pi. \tag{1.73}$$

Below are examples of projectors in matrix representations:

$$\Pi_1 \cong \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Pi_2 \cong \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.74}$$

Π_1 is a rank 1 projector, while Π_2 is rank $2 > 1$. Why we call an operator with the properties in Eq. (1.73) a projector might not be obvious, but the nomenclature is elucidated by the above examples. A projector projects a state into a lower-dimensional subspace of the Hilbert space. Π_1 has the property of taking a three-dimensional vector and projecting it into a 1-dimensional subspace, while Π_2 has the property of taking a three-dimensional vector and projecting it into a 2-dimensional subspace. $\mathbb{I}$ is a projector (though a trivial one), and is a projection from a space to itself. In Fig. 1.6 we visualize the action of Π_1, Π_2 for the case when our vector space is $\mathbb{R}^3$ (but one should keep in mind that this is for the sake of intuition, and the Hilbert spaces we use in quantum mechanics are, of course, complex).

Now let's return to the bra-ket formalism and what projectors look like in this abstract setting. First, recall that we can write an observable A in the form:

$$A = \sum_i a_i |a_i\rangle\langle a_i| \tag{1.75}$$

where $|a_i\rangle$ is the eigenstate of A corresponding to the eigenvalue a_i, and $\{|a_i\rangle\}_i$ is an ONB. In the non-degenerate case, each of the a_is are distinct. However, in general degenerate eigenvalues (where $a_i = a_j$ for some i, j) are possible, and the current form of the expression does not make this particularly clear. With our knowledge of projectors, let us now rewrite the above as:

$$A = \sum_a a\Pi_a \tag{1.76}$$

where each of the as are *distinct* eigenvalues of A, and

$$\Pi_a = \sum_{a_i = a} |a_i\rangle\langle a_i| \tag{1.77}$$

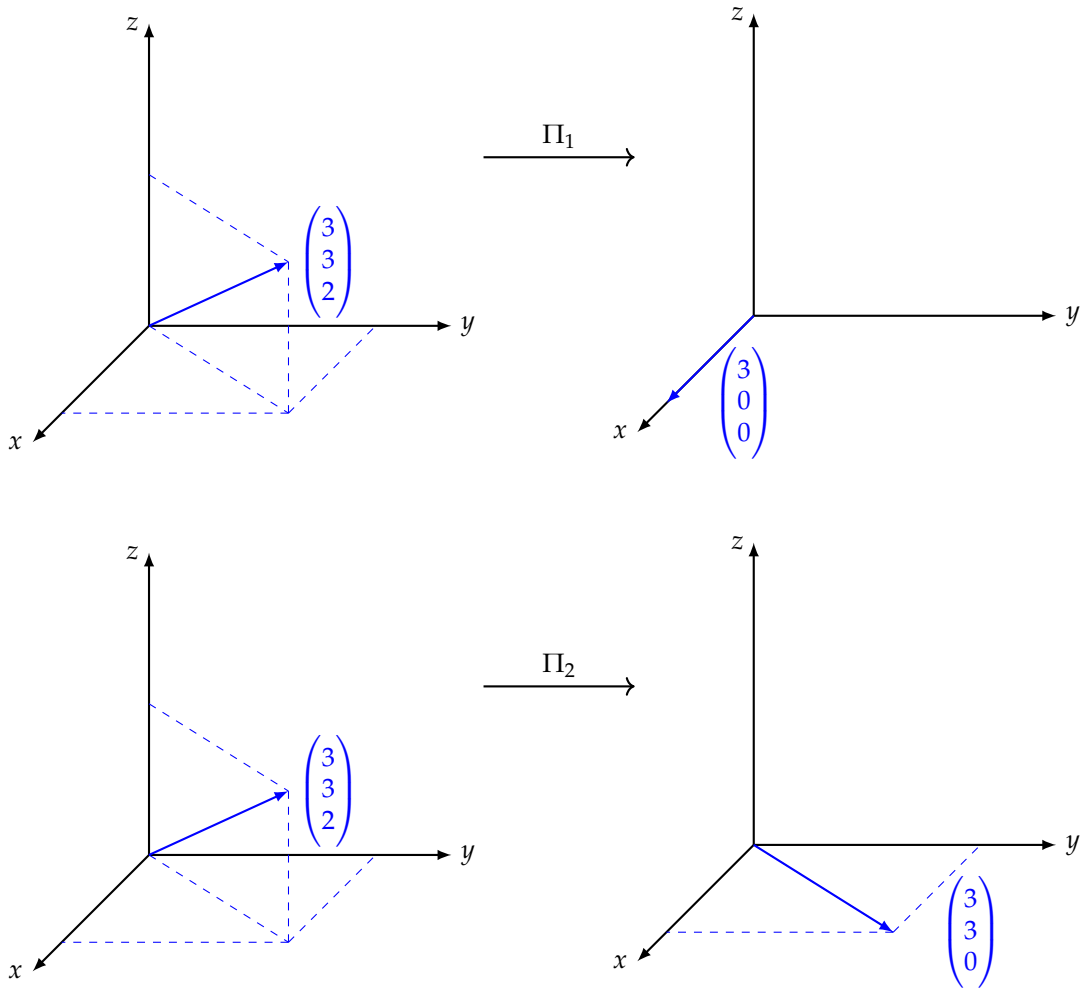

Figure 1.6: Visualization of the action of projectors Π_1, Π_2 (as defined in Eq. (1.74)) on a vector in $\mathbb{R}^3$. Π_1 can be visualized as projecting the given vector onto the one-dimensional subspace that is the x-axis subspace; preserving the x-component of the vector, and nullifying the y and z-components. Π_2 can be visualized as projecting the given vector onto the two-dimensional subspace that is the xy-plane; preserving the x and y components of the vector and nullifying the z-component.

is the *projector onto the eigenvalue-a subspace*. Let us verify that these are indeed projectors. First, we verify that they are Hermitian:

$$\Pi_a^\dagger = \left(\sum_{a_i=a} |a_i\rangle\langle a_i|\right)^\dagger = \sum_{a_i=a} (|a_i\rangle\langle a_i|)^\dagger = \sum_{a_i=a} |a_i\rangle\langle a_i| = \Pi_a. \tag{1.78}$$

where in the second-to-last equality we use that $(|a\rangle\langle b|)^\dagger = |b\rangle\langle a|$ (which follows immediately from the definition of the Hermitian adjoint; the proof is left to the reader!). Next, we show that they are idempotent (that is, they square to themselves):

$$\Pi_a^2 = \left(\sum_{a_i=a} |a_i\rangle\langle a_i|\right)^2 = \sum_{a_i=a}\sum_{a_j=a} |a_i\rangle\langle a_i|a_j\rangle\langle a_j| = \sum_{a_i=a}\sum_{a_j=a} |a_i\rangle\langle a_j|\delta_{ij} = \sum_{a_i=a} |a_i\rangle\langle a_i| = \Pi_a. \tag{1.79}$$

So they are indeed projectors! In this form, we have decomposed the observable A into the parts associated with each eigenvalue in a clear way. These projectors have some properties of note, described in the theorem below.

Proposition

Let $\{\Pi_a\}_a$ be the set of projectors associated to an observable A (with $\Pi_a = \sum_{a_i=a}|a_i\rangle\langle a_i|$ being the projector onto the eigenvalue-a subspace). These projectors are mutually orthogonal:

$$\Pi_i\Pi_j = \delta_{ij}\Pi_i \tag{1.80}$$

and are complete:

$$\sum_a \Pi_a = \mathbb{I}. \tag{1.81}$$

Proof. The idempotency of projectors covers the $i = j$ case in Eq. (1.80), and if $i \neq j$, then the expression is zero as eigenvectors of an observable A corresponding to distinct eigenvalues are orthogonal. The completeness relation is merely a restatement of the resolution of the identity in terms of projectors. □

We are now ready to do our final, complete statement of the axioms of quantum measurement.

Axiom: Quantum measurement (version 3/final)

Let $A = \sum_a a\Pi_a$ be the observable (a Hermitian operator) being measured, where the as are the eigenvalues of A and $\left\{\Pi_a = \sum_{a_i=a}|a_i\rangle\langle a_i|\right\}_a$ are the associated projectors onto the eigenvalue-a subspaces. Let $|\psi\rangle$ be the pre-measurement state.

Dirac postulate: If outcome a is measured, then the post measurement state is given by:

$$|\psi\rangle \mapsto \frac{1}{\sqrt{\langle\psi|\Pi_a|\psi\rangle}}\Pi_a|\psi\rangle. \tag{1.82}$$

Born rule: The probability of measuring outcome a is given by:

$$p(a) = \langle\psi|\Pi_a|\psi\rangle. \tag{1.83}$$

We have now reproduced the form of the Dirac postulate and Born rule shown in the initial table! The above formulation of the measurement axiom(s) is very general, encompassing possible degenerate eigenvalues in the measured observables. Of course, it should be consistent with our previous statement

of the axioms in the case that the eigenvalues are non-degenerate. It is highly recommended that you try this as an exercise first, but we will also give the argument here.

If an eigenvalue a is non degenerate, then $\Pi_a = |a\rangle\langle a|$. Eq. (1.82) then reads:

$$|\psi\rangle \mapsto \frac{1}{\sqrt{\langle\psi|a\rangle\langle a|\psi\rangle}}|a\rangle\langle a|\psi\rangle = \frac{\langle a|\psi\rangle}{\sqrt{|\langle a|\psi\rangle|^2}}|a\rangle = \frac{\langle a|\psi\rangle}{|\langle a|\psi\rangle|}|a\rangle = e^{i\phi}|a\rangle \sim |a\rangle. \tag{1.84}$$

Which is consistent with the previous statement of the Dirac postulate. For the Born rule, we see that Eq. (1.83) reads:

$$p(a) = \langle\psi|a\rangle\langle a|\psi\rangle = |\langle a|\psi\rangle|^2 \tag{1.85}$$

which is again consistent with our previous formulation.

The astute reader may object that we have seemingly ignored the possible complex phase factor sitting out front in Eq. (1.84). However, this is not being sloppy, but instead a fact about quantum states that *global phases are irrelevant*.

Theorem: Irrelevance of global phase

$|\psi\rangle$ and $|\varphi\rangle = e^{i\phi}|\psi\rangle$ correspond to the same physical quantum state.

Proof. The two states are only distinct if we are able to distinguish them in a measurement. However, when calculating the probability of measuring an arbitrary outcome a for any observable A with the Born rule, we find that the two have identical measurement statistics:

$$p_\varphi(a) = \langle\varphi|\Pi_a|\varphi\rangle = \langle\psi|e^{-i\phi}\Pi_a e^{i\phi}|\psi\rangle = \langle\psi|e^{-i\phi}e^{i\phi}\Pi_a|\psi\rangle = \langle\psi|\Pi_a|\psi\rangle = p_\psi(a). \tag{1.86}$$

They therefore represent the same quantum state. □

We however note that *relative* phases *are* significant/relevant. For example, the S_x eigenstates $|+\rangle = \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}$ and $|-\rangle = \frac{|\uparrow\rangle-|\downarrow\rangle}{\sqrt{2}}$ differ by a relative phase, and are hence different quantum states.

Let us return to our motivating example with the spin-1 particle. We established that:

$$S_z^2 = \hbar^2\left(|+\rangle\langle+| + |-\rangle\langle-|\right) \tag{1.87}$$

has a degenerate eigenvalue, with both $|+\rangle$ and $|-\rangle$ being eigenstates with eigenvalue $+\hbar^2$. To deal with this degeneracy, we can use our new projector formalism of measurement. The projector corresponding to the $\hbar^2$ subspace is given by:

$$\Pi_{\hbar^2} = |+\rangle\langle+| + |-\rangle\langle-| \tag{1.88}$$

while the projector corresponding to the eigenvalue 0 subspace is given by:

$$\Pi_0 = |0\rangle\langle 0|. \tag{1.89}$$

So, if we wanted to find the probability of measuring $S_z^2 = \hbar^2$ given a pre-measurement state $|\psi\rangle$, the probability would be given by:

$$p(\hbar^2) = \langle\psi|\Pi_{\hbar^2}|\psi\rangle = |\langle+|\psi\rangle|^2 + |\langle-|\psi\rangle|^2 \tag{1.90}$$

and the post measurement state would be given by:

$$|\psi\rangle \mapsto \frac{1}{\sqrt{\langle\psi|\Pi_{\hbar^2}|\psi\rangle}}\Pi_{\hbar^2}|\psi\rangle = \frac{1}{\sqrt{|\langle+|\psi\rangle|^2 + |\langle-|\psi\rangle|^2}}\left(\langle+|\psi\rangle|+\rangle + \langle-|\psi\rangle|-\rangle\right). \tag{1.91}$$

To conclude this section, we revisit the idea of individual vs. averaged measurement outcomes. We again consider the measurement of an observable $\sum_a a\Pi_a$. As before, the measurement outcomes of individual measurements are given by the eigenvalues a of A. We can again calculate the average outcome/expectation value to be:

$$\langle A\rangle_\psi := \sum_a a p(a) = \sum_a a\langle\psi|\Pi_a|\psi\rangle = \langle\psi|\left(\sum_a a\Pi_a\right)|\psi\rangle = \langle\psi|A|\psi\rangle \tag{1.92}$$

so we see that Eq. (1.48) holds as in the non-degenerate case.

1.5 Compatible and Incompatible Observables

We have seen in the previous sections how quantum measurement is an active process; through measurement the input state is projected into a subspace (corresponding to the measured eigenvalue). This has implications for whether the value of two quantum observables can be simultaneously known; we will illuminate this with a motivating example before exploring this idea more rigorously. Consider a *sequential* Stern-Gerlach experiment, as pictured in Fig. 1.7.

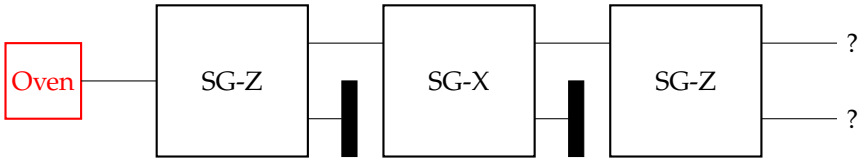

Figure 1.7: Cartoon of a sequential Stern-Gerlach experiment. First, the z-component of the spin of particles coming out of the oven area measured. The $S_z = +\frac{\hbar}{2}$ (spin up/$|\uparrow\rangle$) particles are allowed to go through, while the $S_z = -\frac{\hbar}{2}$ (spin down/$|\downarrow\rangle$) particles are blocked. Then, the x-component of the spin is measured, with the $S_x = +\frac{\hbar}{2}$ (spin right/$|+\rangle$) particles allowed to go through, and the $S_x = -\frac{\hbar}{2}$ (spin left/$|-\rangle$) particles blocked. Finally, the z-component of the spin is measured again.

Let us carry out the calculation to deduce what we would find at the output ports of the above experiment. In the first S_z measurement, there is a 50/50 probability of measuring the spin to be spin-up, and so 50% of the original particles go through. Then recalling that $|\rightarrow\rangle / |\leftarrow\rangle = \frac{|\uparrow\rangle \pm |\downarrow\rangle}{\sqrt{2}}$, we have that $|\uparrow\rangle = \frac{|\rightarrow\rangle + |\leftarrow\rangle}{\sqrt{2}}$ and so we find in the second measurement that $p(\rightarrow) = |\langle\rightarrow|\uparrow\rangle|^2 = \frac{1}{2}$ And $p(\leftarrow) = |\langle\leftarrow|\uparrow\rangle|^2 = \frac{1}{2}$ (we may use the non-degenerate formulation of the Born rule here) so half of the particles go through - one quarter of the original particles from the oven. Now, we note something interesting; when we do the probability calculation for the third measurement, we find $p(\uparrow) = |\langle\uparrow|\rightarrow\rangle|^2 = \frac{1}{2}$ and $p(\downarrow) = |\langle\downarrow|\rightarrow\rangle|^2 = \frac{1}{2}$, so we find that one eighth of the original particles from the oven are present at each of the two output ports. Classically, this is quite strange - a classical measurement does not disturb the system, so when we measure the z-component of spin for the second time, we should expect to get 100% spin-up (as we post-selected on spin-up in the first measurement)! But this is indeed not what happens quantum mechanically - in QM, measurements physically change the state of the system, so there are physical observables that are incompatible with each other (that is, their value cannot be simulataneously precisely known). Let us formalize this notion with a definition:

Definition: Compatible observables

Two observables A, B are *compatible* if:

$$[A, B] := AB - BA = 0. \tag{1.93}$$

The $[\cdot,\cdot]$ appearing in the above definition is known as the *commutator*, and can be interpreted as quantifying the extent to which two objects commute.

Compatible observables are characterized by the following properties:

- The probability $p(a,b)$ for obtaining the outcome a for A and b for B is independent of the order of measurement.
- Compatible measurements do not disturb one another; suppose A is measured with outcome a and then B with outcome b. If A is measured a second time, then the outcome is a again, with certainty.

These properties follow as the result of the following Lemma:

Lemma

Compatible observables have a joint eigenbasis.

Proof. Choose an (orthonormal) eigenbasis $\{|a\rangle\}$ of A. Since $[A,B]=0$, it follows that $\langle a'|[A,B]|a\rangle = 0$ and so:

$$0 = \langle a'|[A,B]|a\rangle = (\langle a'|A)B|a\rangle - \langle a'|B(A|a\rangle) = (a'-a)\langle a'|B|a\rangle \tag{1.94}$$

So if $a \neq a'$, it follows that $\langle a'|B|a\rangle = 0$. If A is non-degenerate, we are immediately done as B is diagonal in the basis $\{|a\rangle\}$ and hence it is also an eigenbasis of B. However, if A is degenerate then $a = a'$ is possible for distinct $|a\rangle, |a'\rangle$. This means that B is block diagonal in the basis $\{|a\rangle\}$. For example in the case where A has two degenerate eigenvalues a_1 and three degenerate eigenvalues a_2, B in this eigenbasis will be block diagonal with two blocks:

$$A = \begin{pmatrix} \begin{matrix} a_1 & 0 \\ 0 & a_1 \end{matrix} & 0 \\ 0 & \begin{matrix} a_2 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_2 \end{matrix} \end{pmatrix} \qquad B = \begin{pmatrix} \blacksquare & 0 \\ 0 & \blacksquare \end{pmatrix} \tag{1.95}$$

Now, we borrow a a known theorem from linear algebra (the Spectral theorem) that a Hermitian matrix can be diagonalized with a unitary matrix - hence, there exists some unitary transformation U (which will have a block-diagonal matrix representation consisting of blocks U_1, U_2 in the same location as the blocks of B) which diagonalizes B:

$$U = \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix} \tag{1.96}$$

Then, we have that $UAU^\dagger = A$ (consider that the blocks of U act on the eigenspaces of A independently, and then cancel to the identity by the condition $UU^\dagger = I$, thus leaving A invariant) and $UBU^\dagger = B'$ is diagonal. In other words, the basis $\left\{U^\dagger|a\rangle\right\}$ is a joint eigenbasis of A, B and we are done. □

Since A, B share an eigenbasis, in this same basis the projectors corresponding to the various are diagonal. It therefore follows that:

$$[\Pi_{A,a}, \Pi_{B,b}] = [\Pi_{A,a}, B] = [A, \Pi_{B,b}] = 0. \tag{1.97}$$

From this fact, we can argue the two properties of compatible observables. First, we show that the probability $p(a,b)$ for obtaining the outcome a for A and b for B is independent of the order of measurement.

Suppose we measure A first, then B. Then the post measurement state after the first measurement is $\frac{\Pi_{a,a}|\psi\rangle}{\sqrt{\langle\psi|\Pi_{A,a}|\psi\rangle}}$, by the Dirac postulate. The probability to measure b in the second measurement (given that we measured a in the first measurement) is then given by the Born rule to be:

$$p(b|a) = \frac{\langle\psi|\Pi_{A,a}\Pi_{B,b}\Pi_{A,a}|\psi\rangle}{\langle\psi|\Pi_{A,a}|\psi\rangle} \tag{1.98}$$

The probability for measuring a and b is then calculated as:

$$\begin{aligned} p(a,b) = p(a\cap b) = p(b|a)p(a) &= \left(\frac{\langle\psi|\Pi_{A,a}\Pi_{B,b}\Pi_{A,a}|\psi\rangle}{\langle\psi|\Pi_{A,a}|\psi\rangle}\right)\langle\psi|\Pi_{A,a}|\psi\rangle \\ &= \langle\psi|\Pi_{A,a}\Pi_{B,b}\Pi_{A,a}|\psi\rangle \\ &= \langle\psi|\Pi_{A,a}^2\Pi_{B,b}|\psi\rangle \\ &= \langle\psi|\Pi_{A,a}\Pi_{B,b}|\psi\rangle \end{aligned} \tag{1.99}$$

where in the second-to-last equality we use that the projectors commute, and in the last equality we use that projectors are idempotent. If we measure B first and then A and go through the same calculation (simply by interchanging $A \leftrightarrow B$), we find the exact same expression, and therefore we conclude that $p(a,b)$ is indepdendent of measurement order.

Let us also demonstrate that the compatible measurements do not disturb one another. If we measure A with outcome a, we have $|\psi\rangle \mapsto |a\rangle$ with $A|a\rangle = a|a\rangle$. If we then measure B, we have $|a\rangle \mapsto \frac{\Pi_{B,b}|a\rangle}{\sqrt{\langle a|\Pi_{B,b}|a\rangle}}$. Using the fact that $A, \Pi_{B,b}$ commute we then observe:

$$A\left(\frac{\Pi_{B,b}|a\rangle}{\sqrt{\langle a|\Pi_{B,b}|a\rangle}}\right) = \frac{\Pi_{B,b}A|a\rangle}{\sqrt{\langle a|\Pi_{B,b}|a\rangle}} = a\left(\frac{\Pi_{B,b}|a\rangle}{\sqrt{\langle a|\Pi_{B,b}|a\rangle}}\right) \tag{1.100}$$

so the post-measurement state after measuring B is still an eigenstate of A, with the same eigenvalue a!

We note an important example of compatible observables. For all observables A, it holds that $[A^n, A^m] = 0$ for all $n, m \in \mathbb{N}$. In other words, all of $\left\{A^k, k \in \mathbb{N}\right\}$ can be simultaneously measured. In the homework, you will be tasked with considering the relationship between the expectation values $|A^k|_\psi = \langle\psi|A^k|\psi\rangle$ of observables A^k, and the measurement probabilities $p(i) = |\langle b_i|\psi\rangle|^2$ where $\left\{|b_i\rangle\right\}_i$ is the eigenbasis of A.

Alongside the discussion of compatible observables runs the consideration of incompatible observables:

Definition: Incompatible observables

Two observables A, B are *incompatible* if:

$$[A,B] \neq 0. \tag{1.101}$$

Incompatible observables have the properties that:

- The probability $p(a,b)$ for obtaining the outcome a for A and b for B is generally dependent of the order of measurement.
- Incompatible measurements disturb one another.

A good example is our motivating example of the S_z and S_x measurements at the beginning of this section. Indeed, $[S_z, S_x] = i\hbar S_y \neq 0$ and so the two observables are incompatible.

1.6 Position and Momentum

In this section, we discuss two key examples of observables - namely position and momentum. We discuss their eigenstates, which take on a continuous range. We discuss their commutation relation which is the axiomatic foundation of how they relate. We derive the momentum operator in the position basis, and show that position and momentum wavefunctions relate via a Fourier transform. Finally, we will derive the famous Heisenberg uncertainty relation.

The spin observables we have considered so far in this course have taken on discrete values (e.g. the eigenvalues of S_z for a spin-1/2 particle are $\pm\frac{\hbar}{2}$). However, the eigenvalues of position and momentum are continuous; for example, a (free) particle in one dimension can have $x \in (-\infty, \infty)$. To handle this, we will content ourselves to generalize the finite-dimensional formalism to the infinite-dimensional setting in the natural way, sweeping the functional analytic details under the rug.

So, we consider the (one-dimensional) position operator X, with a complete and orthonormal set of eigenstates $\{|x\rangle\}$ that satisfy $X|x\rangle = x|x\rangle$ with x the position eigenvalue. The resolution of identity (as we discussed in a previous section) reads $\int dx|x\rangle\langle x| = \mathbb{I}$, and the orthonormality relation reads $\langle x'|x\rangle = \delta(x - x')$.

We can also consider the (one-dimensional) linear momentum operator P, which has a complete and orthonormal set of eigenstates $\{|p\rangle\}$ that satisfy $P|p\rangle = p|p\rangle$ with p the momentum eigenvalue. The resolution of identity reads $\int dp|p\rangle\langle p| = \mathbb{I}$ and the orthonormality relation reads $\langle p'|p\rangle = \delta(p - p')$.

We define the position and momentum wavefunctions $\psi(x)/\tilde{\psi}(p)$ as the expansion coefficients of a state $|\psi\rangle$ in the position/momentum bases:

$$|\psi\rangle = \left(\int dx|x\rangle\langle x|\right)|\psi\rangle = \int dx\langle x|\psi\rangle|x\rangle = \int dx\psi(x)|x\rangle \tag{1.102}$$

$$|\psi\rangle = \left(\int dp|p\rangle\langle p|\right)|\psi\rangle = \int dp\langle p|\psi\rangle|p\rangle = \int dp\tilde{\psi}(p)|p\rangle. \tag{1.103}$$

Having introduced some operators with continuous spectra, let us return briefly back to our example of higher moments, namely expectation values of X^n. Calculating the expectation of X, we have:

$$\langle X\rangle_\psi = \langle\psi|X|\psi\rangle = \int dxdy\langle\psi|x\rangle\langle x|X|y\rangle\langle y|\psi\rangle = \int dxdy\psi^*(x)y\delta(x-y)\psi(y) = \int dxx|\psi(x)|^2 \tag{1.104}$$

which is the average position of the state $|\psi\rangle$. Calculating the expectation of X^2, we have:

$$\left\langle X^2\right\rangle_\psi = \langle\psi|X|\psi\rangle = \int x^2|\psi(x)|^2 \tag{1.105}$$

which gives information about the width - $|\psi(x)|^2$ has the interpretation of a probability density in position space, and the expectation values of X^n gives information about the distribution. For example the variance yields a measure of the deviation from the mean value:

$$\left\langle(\Delta X)^2\right\rangle_\psi = \left\langle(X - I\langle X\rangle)^2\right\rangle_\psi = \left\langle X^2\right\rangle_\psi - \langle X\rangle_\psi^2 \tag{1.106}$$

We can also calculate the skewedness $\left\langle(\Delta X)^3\right\rangle$, which is a measure of how asymmetric the distribution is, and so on. All moments taken together contain the same information[14] as the probability density $|\psi(x)|^2$.

Having now a better understanding of moments and continuous probability densities, let us discuss the fundamental relationship between position and momentum.

[14] Well, not quite, there are some pathological counterexamples

Axiom: Canonical commutation relations

The commutation relations between position and momentum are:

$$[X_i, X_j] = 0, \quad [P_i, P_j] = 0, \quad [X_i, P_j] = i\hbar\delta_{ij}\mathbb{I} \tag{1.107}$$

where the subscript denotes the component of position/momentum, and δ_{ij} is the Kronecker delta.

Note that there is a noteworthy relationship between classical and quantum mechanics:

$$[\cdot,\cdot]_{\text{classical}} \mapsto \frac{[\cdot,\cdot]}{i\hbar} \tag{1.108}$$

where $[\cdot,\cdot]_{\text{classical}}$ denotes the Poisson bracket.

We can understand the notion of momentum as a generator of translations using these commutation relations. To this end, let us define:

Definition: Translation operator

The translation operator is defined as the following imaginary exponential:

$$\mathcal{T}(\Delta x) := e^{-i\frac{P}{\hbar}\Delta x} \tag{1.109}$$

where P is the momentum operator and Δx some real number.

We now claim that the translation operator lives up to its namesake:

Proposition

The translation operator translates a position eigenket, that is:

$$\mathcal{T}(\Delta x)|x\rangle = |x + \Delta x\rangle \tag{1.110}$$

Proof. Let us Taylor expand the translation operator (recalling the Taylor expansion of the exponential - of course, this is how exponentials of operators are formally defined):

$$\mathcal{T}(\Delta x) = \mathbb{I} - i\frac{P}{\hbar}\Delta x + O((\Delta x)^2). \tag{1.111}$$

Furthermore, let us recall the eigenvalue equation $X|x\rangle = x|x\rangle$. Now, we consider $\mathcal{T}(\Delta x)|x\rangle$:

$$\begin{aligned} X\mathcal{T}(\Delta x)|x\rangle &= X\left(\mathbb{I} - i\frac{P}{\hbar}\Delta x + O((\Delta x)^2)\right)|x\rangle \\ &= X|x\rangle - i\frac{\Delta x}{\hbar}XP|x\rangle + O((\Delta x)^2) \\ &= x|x\rangle - i\frac{\Delta x}{\hbar}\left(i\hbar\mathbb{I} + PX\right)|x\rangle + O((\Delta x)^2) \\ &= x|x\rangle + \Delta x|x\rangle - i\frac{\Delta x}{\hbar}Px|x\rangle + O((\Delta x)^2) \\ &= (x + \Delta x)\left(\mathbb{I} - i\frac{P}{\hbar}\Delta x\right)|x\rangle + O((\Delta x)^2) \\ &= (x + \Delta x)\mathcal{T}(\Delta x)|x\rangle + O((\Delta x)^2) \end{aligned} \tag{1.112}$$

where in the third equality we note the application of the canonical commutation relation. Also note in the fifth equality we introduce a term $\propto (\Delta x)^2$ but we may do this freely as such terms are absorbed by $O((\Delta x)^2)$. The conclusion is (up to terms of order $(\Delta x)^2$) that $\mathcal{T}(\Delta x)|x\rangle$ is an eigenket of position with eigenvalue $x + \Delta x$ and so $\mathcal{T}(\Delta x)|x\rangle = |x + \Delta x\rangle$. So, for infinitesimal translations (for which we can neglect terms $O((\Delta x)^2)$) the claim holds. We are then able to conclude the result for general translations by considering a general translation as a composition of many infinitesimal ones. □

Note: We could instead have posited as an axiom the above property of the translation operator/momentum as a generator of translations, and from there derived the canonical commutation relations (and in fact Sakurai derives things in this order) - there is not a huge difference as one set of relations can always be derived from the other.

Using the translation operator, we can derive the form of the momentum operator in the position basis - this will be useful for deriving the form of the Schrodinger equation in the position basis from the general basis-independent equation, as we will do in the next part of the course.

Proposition: Momentum operator in position basis

For a general state $|\alpha\rangle$, we have:

$$\langle x|P|\alpha\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|\alpha\rangle \tag{1.113}$$

Proof. We consider operating the translation operator $\mathcal{T}(\Delta x)$ on the general state $|\alpha\rangle$. We consider small Δx so as to neglect terms of order $(\Delta x)^2$ and higher:

$$\begin{aligned}
\left(\mathbb{I} - i\frac{P}{\hbar}\Delta x\right)|\alpha\rangle &= \mathcal{T}(\Delta x)|\alpha\rangle \\
&= \int dx' \mathcal{T}(\Delta x)|x'\rangle\langle x'|\alpha\rangle \\
&= \int dx' |x' + \Delta x\rangle\langle x'|\alpha\rangle \\
&= \int dx' |x'\rangle\langle x' - \Delta x|\alpha\rangle \\
&= \int dx' |x'\rangle \left(\langle x'|\alpha\rangle - \Delta x' \frac{\partial}{\partial x'} \langle x'|\alpha\rangle\right)
\end{aligned} \tag{1.114}$$

where in the second equality we have inserted the resolution of the identity, in the third equality we use the translation property derived above, in the fourth equality we make the substitution $x' \to x' - \Delta x$, and in the fifth equality we Taylor expand $\langle x' - \Delta x|\alpha\rangle$. Now, if we equate order Δx terms on both sides, we obtain:

$$-i\frac{P}{\hbar}\Delta x|\alpha\rangle = \int dx' |x'\rangle(-\Delta x' \frac{\partial}{\partial x'} \langle x'|\alpha\rangle) \tag{1.115}$$

Multiplying both sides by $i\hbar\langle x|$ and cancelling out the Δxs:

$$\langle x|P|\alpha\rangle = \int dx' \langle x|x'\rangle \left(-i\hbar \frac{\partial}{\partial x'} \langle x'|\alpha\rangle\right) = \int dx' \delta(x - x') \left(-i\hbar \frac{\partial}{\partial x'} \langle x'|\alpha\rangle\right) \tag{1.116}$$

the delta function picks out $x' = x$ in the integral and so we obtain:

$$\langle x|P|\alpha\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|\alpha\rangle \tag{1.117}$$

as claimed. □

With this representation of the momentum operator derived, we are now able to show that the position and momentum space wavefunctions (Eqs. (1.102), (1.103)) are related via a Fourier transform.

Proposition: Fourier transform between position and momentum

Given a state $|\psi\rangle$, its momentum wavefunction $\tilde{\psi}(p)$ and position wavefunction $\psi(x)$ are related by a Fourier transform, that is:

$$\tilde{\psi}(p) \propto \int dx e^{-ipx/\hbar}\psi(x). \tag{1.118}$$

Proof. Recall that:

$$\tilde{\psi}(p) = \langle p|\psi\rangle = \int dx \langle p|x\rangle\langle x|\psi\rangle = \int dx \langle p|x\rangle\psi(x). \tag{1.119}$$

by definition of the momentum/position wavefunctions and using the resolution of the identity. So, it suffices to show that $\langle p|x\rangle \propto e^{-ipx/\hbar}$.

We recall the representation $\langle x|P|\alpha\rangle = -i\hbar\dfrac{\partial}{\partial x}\langle x|\alpha\rangle$ that we derived just before. Now, set $|\alpha\rangle = |p\rangle$ a momentum eigenstate. We then have that:

$$\langle x|P|p\rangle = \langle x|p|p\rangle = p\langle x|p\rangle \tag{1.120}$$

using the eigenvalue equation, and also that:

$$\langle x|P|p\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|p\rangle \tag{1.121}$$

using our previous result. We therefore obtain the equation:

$$p\langle x|p\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|p\rangle \tag{1.122}$$

which is a first-order ODE which is easily solved:

$$\langle x|p\rangle \propto e^{ipx/\hbar} \tag{1.123}$$

and therefore:

$$\langle p|x\rangle = \langle x|p\rangle^* \propto e^{-ipx/\hbar} \tag{1.124}$$

and so we are done. □

There is one detail we skipped in the above proof; namely, the normalization of $\langle x|p\rangle$. If we want to determine N in $\langle x|p\rangle = Ne^{ipx/\hbar}$, we can consider:

$$\begin{aligned}\delta(x-x') = \langle x|x'\rangle &= \int dp\langle x|p\rangle\langle p|x'\rangle \\ &= |N|^2\int dp e^{ip(x-x')/\hbar} \\ &= |N|^2 2\pi\hbar\delta(x-x')\end{aligned} \tag{1.125}$$

where the last integral is a standard one in Fourier Analysis. comparing the two sides of the equation, we can conclude:

$$N = \frac{1}{\sqrt{2\pi\hbar}}. \tag{1.126}$$

Now, we move to a discussion of uncertainty relations - the most famous of which is the Heisenberg uncertainty relation of position and momentum. This relation has key physical implications - for example

being responsible for the size scale of atoms. Let us explore the HUP with an example before going into its formal derivation.

We consider a variation on the sequential Stern-Gerlach experiment involving position and momentum; in other words, a single-slit diffraction experiment[15]!

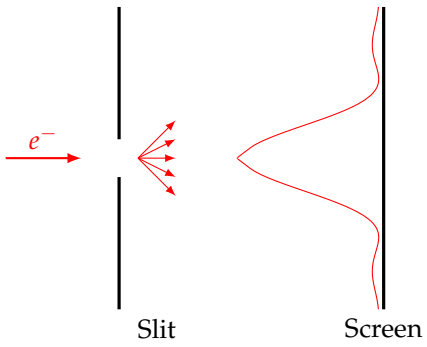

Figure 1.8: Cartoon of a single-slit diffraction experiment. A beam of electrons is fired through a slit, which amounts to a position measurement of the electrons at the slit. The electrons then propogate until they hit the screen, which results in a diffraction pattern (after many electrons are fired) - the momentum of the electron at the slit can then be inferred from the position which the electron hits the screen.

To understand this experiment quantum-mechanically, we consider that the position of electrons is measured as they pass through a slit; at this point, a reasonable description of the position wavefunction of the electrons is the uniform wavefunction extending across the slit (which we assume to have width $2d$):

$$\psi(x) = \begin{cases} \frac{1}{\sqrt{2d}} & -d \leq x \leq d \\ 0 & \text{otherwise} \end{cases} \tag{1.127}$$

We can then take the Fourier transform to find the momentum wavefunction:

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi(x) e^{-ipx/\hbar} = \sqrt{\frac{\hbar}{\pi d}} \frac{\sin(\frac{pd}{\hbar})}{p} \tag{1.128}$$

(we skip the algebra here; but this is one special case where one can easily take the Fourier transform. Try it!) The modulus square of this momentum wavefunction yields the diffraction pattern that we see on the experiment screen.

This experiment is interesting to us here because it shows us how both position and momentum cannot be precisely known; Looking at the expression for $\tilde{\psi}(p)$ above, we see that if we decrease the slit size d (i.e. we are more certain of the position of the electron) then the momentum wavefunction spreads out and we are less certain of it (and vise versa) - see Fig. 1.9 for a graphical sketch of this.

Let us now derive the HUP. To begin, we will prove a more general uncertainty principle between operators. To this end, we will require a Lemma concerning inner products of vectors:

Lemma: Cauchy-Shwartz inequality

Let $|\alpha\rangle, |\beta\rangle$ be two vectors. Then,

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2. \tag{1.129}$$

[15]This may be familiar to you from a homework assignment in PHYS 200 - if not, we go through it again briefly here together.

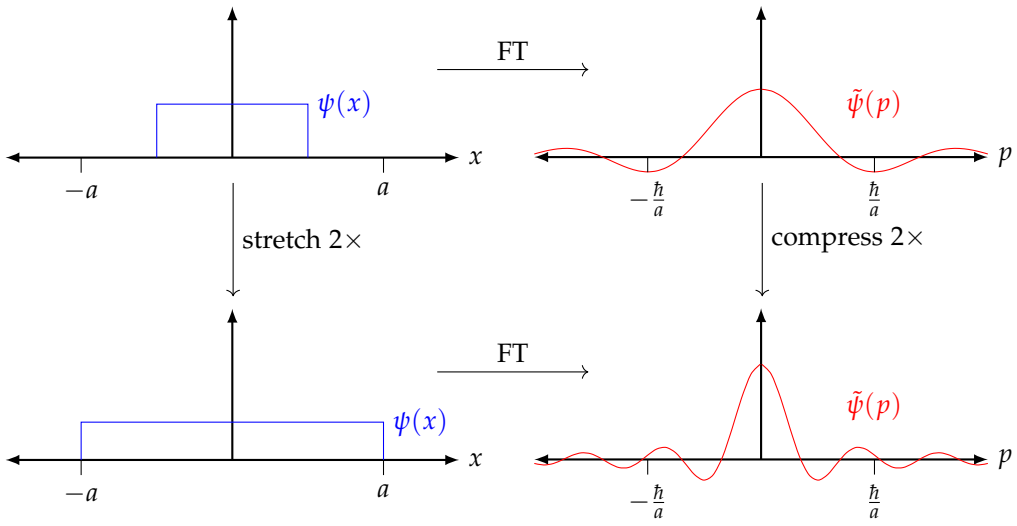

Figure 1.9: Graphical Depiction of the impact of varying the width of the position wavefunction (slit size) on the momentum wavefunction (diffraction pattern width). As we stretch out the position wavefunction (and the position of the electron becomes more uncertain), the momentum wavefunction compresses, and we are more certain about the momentum of the electron at the slit. However, the position and momentum uncertainty cannot be decreased simultaneously - something that is rigorously characterized in the Heisenberg Uncertainty Relation.

Proof. Since $\langle v|v\rangle \geq 0$ for any vector $|v\rangle$, it follows that:

$$(\langle\alpha| + \lambda^*\langle\beta|)(|\alpha\rangle + \lambda|\beta\rangle) \geq 0 \tag{1.130}$$

for all $\lambda \in \mathbb{C}$. Now, set $\lambda = -\frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}$. Multiplying out the resulting expression, we obtain:

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle - |\langle\alpha|\beta\rangle|^2 \geq 0 \tag{1.131}$$

which proves the claim. □

With this Lemma under our belt, we have the tools in place to prove the following:

Theorem: General uncertainty relation

Let A, B be two observables, and define $\Delta A := A - \langle A\rangle_\psi \mathbb{I}$, $\Delta B := B - \langle B\rangle_\psi \mathbb{I}$, (where the expectation value is taken with respect to some state $|\psi\rangle$). It then follows that:

$$\left\langle (\Delta A)^2 \right\rangle_\psi \left\langle (\Delta B)^2 \right\rangle_\psi \geq \frac{1}{4}|\langle [A,B]\rangle_\psi|^2. \tag{1.132}$$

Proof. Let $|\alpha\rangle = (\Delta A)|\psi\rangle$ and $|\beta\rangle = (\Delta B)|\psi\rangle$. Applying Cauchy-Shwartz, we obtain:

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2 \implies \left\langle (\Delta A)^2 \right\rangle_\psi \left\langle (\Delta B)^2 \right\rangle_\psi \geq |\langle \Delta A \Delta B\rangle_\psi|^2. \tag{1.133}$$

Now, notice that we can decompose:

$$\Delta A \Delta B = \frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\} \tag{1.134}$$

Where $\{\Delta A, \Delta B\} = \Delta A \Delta B + \Delta B \Delta A$ denotes the *anti-commutator* of ΔA and ΔB. Now, we claim that $\langle\psi|[\Delta A, \Delta B]|\psi\rangle$ is imaginary. To see this, note:

$$[\Delta A, \Delta B]^\dagger = (\Delta A \Delta B - \Delta B \Delta A)^\dagger = \Delta B \Delta A - \Delta A \Delta B = -[\Delta A, \Delta B] \tag{1.135}$$

where we have used that $\Delta A, \Delta B$ are Hermitian, and that $(AB)^\dagger = B^\dagger A^\dagger$ (this can be proven easily from the definition of the adjoint using the dual correspondence - it is a good exercise to check it)! From the above equation, we obtain:

$$(\langle\psi|[\Delta A, \Delta B]|\psi\rangle)^* = -\langle\psi|[\Delta A, \Delta B]|\psi\rangle \tag{1.136}$$

from which we would conclude that $\langle\psi|[\Delta A, \Delta B]|\psi\rangle$ must be purely imaginary.

We can also verify that $\langle\psi|\{\Delta A, \Delta B\}|\psi\rangle$ is real (check)! The upshot is then that we can write:

$$\langle \Delta A \Delta B \rangle_\psi = \frac{1}{2} \left\langle [\Delta A, \Delta B] \right\rangle_\psi + \frac{1}{2} \left\langle \{\Delta A, \Delta B\} \right\rangle_\psi \tag{1.137}$$

with the first term purely imaginary, and the second term purely real. Then, using that for a complex number $z = a + bi$ that the modulus square is $|z|^2 = |a|^2 + |b|^2$, we find:

$$|\langle \Delta A \Delta B \rangle_\psi|^2 = \frac{1}{4} |\left\langle [\Delta A, \Delta B] \right\rangle_\psi|^2 + \frac{1}{4} |\left\langle \{\Delta A, \Delta B\} \right\rangle_\psi|^2 \geq \frac{1}{4} |\left\langle [\Delta A, \Delta B] \right\rangle_\psi|^2 \tag{1.138}$$

where the last inequality follows from both terms being positive. Combining this result with (1.133), we find:

$$\left\langle (\Delta A)^2 \right\rangle_\psi \left\langle (\Delta B)^2 \right\rangle_\psi \geq \frac{1}{4} |\left\langle [\Delta A, \Delta B] \right\rangle_\psi|^2 \tag{1.139}$$

Finally, note that since the identity commutes with all operators that:

$$[\Delta A, \Delta B] = [A - \langle A \rangle_\psi \mathbb{I}, B - \langle B \rangle_\psi \mathbb{I}] = [A, B] \tag{1.140}$$

and so:

$$\left\langle (\Delta A)^2 \right\rangle_\psi \left\langle (\Delta B)^2 \right\rangle_\psi \geq \frac{1}{4} |\left\langle [A, B] \right\rangle_\psi|^2 \tag{1.141}$$

which is exactly what we wished to show. □

Corollary: Heisenberg uncertainty principle

The position operator X and momentum operator P (in the same direction) satisfy:

$$\left\langle (\Delta X)^2 \right\rangle_\psi \left\langle (\Delta P)^2 \right\rangle_\psi \geq \frac{\hbar^2}{4} \tag{1.142}$$

Proof. This follows immediately by setting $A = X, B = P$ in the general uncertainty relation, and applying the canonical commutation relation $[X, P] = i\hbar\mathbb{I}$. □

Having now proven the Heisenberg uncertainty relation, it may be of interest to ask which wavefunctions saturate it; that is, for which $|\psi\rangle$ does it follow that:

$$\left\langle (\Delta X)^2 \right\rangle_\psi \left\langle (\Delta P)^2 \right\rangle_\psi = \frac{\hbar^2}{4}? \tag{1.143}$$

Proposition: Gaussian wavepackets saturate the HUP

Gaussian Wavepackets, that is, states with position wavefunctions:

$$\langle x|\psi\rangle = (2\pi d^2)^{-1/4} e^{-\frac{x^2}{4d^2}} \tag{1.144}$$

saturate the Heisenberg uncertainty principle.

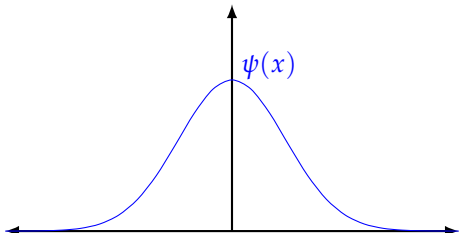

Figure 1.10: Plot of a Gaussian wavefunction (with mean zero).

Proof. First, note that:

$$\left\langle (\Delta A)^2 \right\rangle = \left\langle (A - \langle A\rangle \mathbb{I})^2 \right\rangle = \left\langle A^2 - 2\langle A\rangle A + \langle A\rangle^2 \mathbb{I} \right\rangle = \left\langle A^2 \right\rangle - 2\langle A\rangle^2 + \langle A\rangle^2 = \left\langle A^2 \right\rangle - \langle A\rangle^2 \tag{1.145}$$

So we calculate $\langle A\rangle_\psi$ and $\left\langle A^2 \right\rangle_\psi$ for $A = X, P$. First, note that $\langle X\rangle_\psi = 0$ as the wavefunction is symmetric about $x = 0$ ($\psi(x) = \psi(-x)$). Therefore:

$$\left\langle (\Delta X)^2 \right\rangle_\psi = \left\langle X^2 \right\rangle_\psi = \int_{-\infty}^{\infty} dx x^2 \left((2\pi d^2)^{-1/4} e^{-\frac{x^2}{4d^2}} \right)^2 = d^2 \tag{1.146}$$

Since $\dfrac{\mathrm{d}\psi(x)}{\mathrm{d}x}$ is anti-symmetric, $\langle P\rangle_\psi = 0$ and hence:

$$\left\langle (\Delta P)^2 \right\rangle_\psi = \left\langle P^2 \right\rangle_\psi = \int_{-\infty}^{\infty} dx \left((2\pi d^2)^{-1/4} e^{-\frac{x^2}{4d^2}} \right) \left(-i\hbar \frac{\mathrm{d}}{\mathrm{d}x} \right)^2 \left((2\pi d^2)^{-1/4} e^{-\frac{x^2}{4d^2}} \right) = \frac{\hbar^2}{4d^2} \tag{1.147}$$

where we have used the representation of the momentum operator in the position basis. We therefore conclude for Gaussian wavepackets that:

$$\left\langle (\Delta X)^2 \right\rangle_\psi \left\langle (\Delta P)^2 \right\rangle_\psi = \frac{\hbar^2}{4} \tag{1.148}$$

and hence the claim is proven. □

Note that the above argument generalizes easily to the case where $\langle X\rangle_\psi \neq 0$ for the Gaussian wavepacket (i.e. a Gaussian shifted from the origin). Actually, the relationship between Gaussian wavepackets and HUP saturation is even stronger.

Proposition: HUP Saturation $\iff$ Gaussian Wavpackets

Gaussian wavepackets are the only wavefunctions which saturate the Heisenberg uncertainty principle.

Proof. We have already shown Gaussian wavepackets $\implies$ HUP saturation in the previous proof. So, what is left to do is show that if the HUP is saturated, then the state must be a Gaussian wavepacket.

First, we go back to our derivation of the general uncertainty principle, where we invoked the Cauchy-Shwartz inequality. It can be shown that the CS-inequality is saturated with equality if and only if $|\alpha\rangle = c|\beta\rangle$ for some $c \in \mathbb{C}$ (i.e. the two vectors in question are linearly independent). Therefore, in the context of position and momentum, we require:

$$(\Delta X)|\psi\rangle = c(\Delta P)|\psi\rangle \tag{1.149}$$

for some complex c. Further, in our derivation we threw away the anti-commutator term; for minimum uncertainty we require that this term be zero, i.e. $|\langle\{\Delta X, \Delta P\}\rangle|^2 = 0$. But we established in the uncertainty relation argument that this was real, so if it is zero, it follows that c appearing in Eq. (1.149) is purely imaginary. Theerofre if we consider the ODE defined by Eq. (1.149) in the position basis, we obtain:

$$x\psi(x) = ci\hbar \frac{\partial}{\partial x}\psi(x) \tag{1.150}$$

where we have assumed $\langle X\rangle_\psi = \langle P\rangle_\psi = 0$ for simplicity (however this assumption can be relaxed - $\langle X\rangle$ will only shift the Gaussian some amount from the origin, and $\langle P\rangle$ will only tack on a physically irrelevant phase factor - check this if you like!). It can then be easily checked that (for imaginary c) that the solutions to the above ODE are Gaussians, which completes the proof of the claim. □

2 Quantum Dynamics

2.1 Introduction to the Schrödinger Equation

Recall Table 1 which laid out the axioms of quantum mechanics; up until this point we have discussed quantum states, as well as quantum measurement (projective dynamics) - we will now discuss the last entry in the table, which concerns the unitary dynamics/evolution of quantum states.

Axiom: The Schrödinger Equation

The evolution of a quantum state $|\psi(t)\rangle$ evolving under the influence of a Hamiltonian operator H (which describes the energy of the system) is given by the Schrödinger Equation:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H|\psi(t)\rangle \tag{2.1}$$

Many quantum phenomena can be derived from the Schrödinger equation; for example the shape of orbitals and the energy spectra of the Hydrogen atom, which is a system you (hopefully) solved in a previous course.

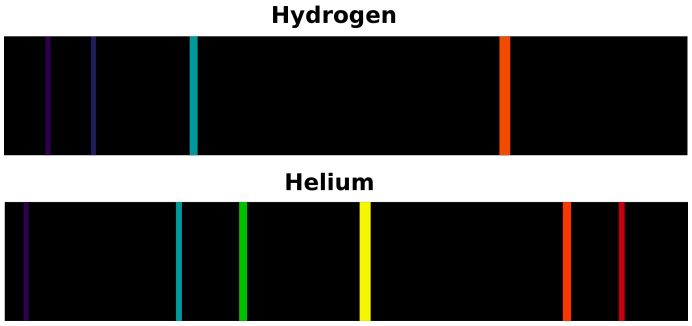

Figure 2.1: Emission spectrum of Hydrogen and Helium, which can be derived by analyzing the Schrödinger equation (though we note that the Hydrogen atom is one of the (very few) systems which are analytically solvable, and the Helium atom must be treated by approximation techniques). Image created by Ranjithsiji, licensed under the Creative Commons Attribution-Share Alike 4.0 International license.

Above, we give the basis-independent form of the Schrödinger equation. This differs from the form of the SE that you might be more familiar with, which likely takes the form as in Eq. (1.7). This form is the SE written in the position basis; we will begin by deriving this form (in 1-D) from the general/basis-independent equation.

In general, the energy (Hamiltonian) operator H has two terms, a (position-dependent) potential term and a kinetic term:

$$H = \underbrace{V(X)}_{\text{potential}} + \underbrace{\frac{P^2}{2m}}_{\text{kinetic}} \tag{2.2}$$

where X, P are the position/momentum operators and m is the mass of the particle. Some familiar forms of V you may have encountered in previous courses are $V(X) = 0$ for the free particle, $V(X) = \frac{1}{2}m\omega^2 X^2$ for the quantum harmonic oscillator (which we will discuss later in this chapter), and $V(R) = -\frac{e^2}{4\pi\epsilon_0}\frac{1}{R}$ for the Hydrogen atom.

Let us now express this basis-independent expression of the Hamiltonian in the position basis.

Proposition: Schrödinger equation in position basis

In the position basis, the Schrödinger equation with Hamiltonian Eq. (2.2) reads:

$$i\hbar \frac{\partial}{\partial t} \psi(x,t) = \left[V(x,t) - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right] \psi(x,t) \tag{2.3}$$

Proof. Starting with the potential term, we have:

$$\begin{aligned} V(X) &= \left(\int dx |x\rangle\langle x| \right) V(X) \left(\int dx' |x'\rangle\langle x'| \right) \\ &= \int dx dx' |x\rangle \left(\langle x|V(X)|x'\rangle \right) \langle x'| \\ &= \int dx dx' |x\rangle V(x') \langle x|x'\rangle |x'\rangle \\ &= \int dx dx' |x\rangle V(x') \delta(x-x') |x'\rangle \\ &= \int dx |x\rangle V(x) \langle x| \end{aligned} \tag{2.4}$$

where we have inserted the resolution of the identity twice in the first line, used that $V(X)|x'\rangle = V(x')|x'\rangle$ (eigenvalue equation for position) in the third line, and used the orthonormality of the position basis in the fourth line.

For the kinetic term, we have:

$$\begin{aligned} P^2|\psi\rangle &= \left(\int dx |x\rangle\langle x| \right) P \left(\int dx' |x'\rangle\langle x'| \right) P|\psi\rangle \\ &= \int dx dx' |x\rangle \left(\langle x|P|x'\rangle \right) \left(\langle x'|P|\psi\rangle \right) \\ &= (-i\hbar)^2 \int dx dx' |x\rangle \frac{\partial}{\partial x} \langle x|x'\rangle \frac{\partial}{\partial x'} \frac{\partial x'}{\partial \psi} \\ &= -\hbar^2 \int dx dx' |x\rangle \frac{\partial}{\partial x} \delta(x-x') \frac{\partial}{\partial x'} \langle x'|\psi\rangle \\ &= -\hbar^2 \int dx |x\rangle \frac{\partial}{\partial x} \frac{\partial}{\partial x} \langle x|\psi\rangle \\ &= -\hbar^2 \int dx |x\rangle \frac{\partial^2}{\partial x^2} \langle x|\psi\rangle \end{aligned} \tag{2.5}$$

Where in the first equality we insert the resolution of the identity twice, in the third equality we used the previously derived expression for momentum in the position basis (Eq. (1.117)), in the fourth equality we use the orthonormality of the position basis, and we use the subsequent dirac delta to carry out the integrals. Now if we multiply by a position eigenbra on both sides, we have:

$$\begin{aligned} \langle x|P^2|\psi\rangle &= \langle x| \left(-\hbar^2 \int dx' |x'\rangle \frac{\partial^2}{\partial x'^2} \langle x'|\psi\rangle \right) \\ &= -\hbar^2 \int dx' \langle x|x'\rangle \frac{\partial^2}{\partial x'^2} \langle x'|\psi\rangle \\ &= -\hbar^2 \int dx' \delta(x-x') \frac{\partial^2}{\partial x'^2} \langle x'|\psi\rangle \\ &= -\hbar^2 \frac{\partial^2}{\partial x^2} \langle x|\psi\rangle \end{aligned} \tag{2.6}$$

We conclude that:

$$\langle x|P^2|\psi\rangle = -\hbar \frac{\partial^2}{\partial x^2} \langle x|\psi\rangle \tag{2.7}$$

Putting it all together; let us multiply the basis-independent SE (Eq. (2.1)) on the right by a position eigenbra:

$$\langle x|i\hbar \frac{\partial}{\partial t} |\psi\rangle = \langle x|H|\psi\rangle = \langle x|V(x)|\psi\rangle + \langle x|\frac{P^2}{2m}|\psi\rangle \tag{2.8}$$

The prior results then yield:

$$i\hbar \frac{\partial}{\partial t} \langle x|\psi\rangle = V(x)\langle x|\psi\rangle - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \langle x|\psi\rangle \tag{2.9}$$

and since $\langle x|\psi\rangle$ is just the position wavefunction, we have now successfully derived the (familiar) 1-D SE in the position basis (also, let us now insert the time parameter into the wavefunction and the potential, as the wavefunction is time-dependent and the potential also can be time-dependent depending on the system; e.g. a time-varying magnetic field):

$$i\hbar \frac{\partial}{\partial t} \psi(x,t) = \left[V(x,t) - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right] \psi(x,t) \tag{2.10}$$

□

In 3D we have the generalization:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r},t) = \left[V(\mathbf{r},t) - \frac{\hbar^2}{2m} \nabla^2 \right] \psi(\mathbf{r},t) \tag{2.11}$$

where:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2.12}$$

is the Laplacian.

2.2 Spin Precession and Medical Imaging

Let us explore a concrete example of Schrödinger evolution of a quantum system. Namely, we analyze how a spin-1/2 system evolves in a constant magnetic field.

Classically, the energy of a magnetic dipole with dipole moment $\mathbf{m}$ in a magnetic field $\mathbf{B}$ is given by:

$$E = -\mathbf{m} \cdot \mathbf{B} \tag{2.13}$$

with the minus sign present as it is energetically favourable for the dipole to align with the magnetic field. In analogy, the quantum mechanical Hamiltonian for the system is given by:

$$H = -\frac{e}{m_e c} \mathbf{S} \cdot \mathbf{B} \tag{2.14}$$

with e the charge of the electron, m_e the mass of the electron, c the speed of light, and $\mathbf{S} = (S_x, S_y, S_z)^T$ a vector of spin operators. Let us suppose that we have a constant field aligned in the $\hat{\mathbf{z}}$ direction, so $\mathbf{B} = B\hat{\mathbf{z}}$. Then the above reduces to:

$$H = -\frac{eB}{m_e c} S_z \tag{2.15}$$

Defining $\omega = \frac{|e|B}{m_e c}$, and recalling that the electron charge is negative, the Hamiltonian can be written as:

$$H = \omega S_z \tag{2.16}$$

The choice of notation ω, traditionally used to represent some frequency, will become clear shortly.

We now look at how an electron spin, originally prepared in the spin state $|\psi(0)\rangle = |\rightarrow\rangle = \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}$, evolves under the influence of this Hamiltonian; that is, we wish to solve the time-dependent SE:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \omega S_z |\psi(t)\rangle \tag{2.17}$$

First, we recall the eigenstates and eigenvalues of S_z, namely that $S_z|\uparrow / \downarrow\rangle = \pm\frac{\hbar}{2}|\uparrow / \downarrow\rangle$ and hence taking the dual $\langle\uparrow / \downarrow|S_z = \pm\langle\uparrow / \downarrow|\frac{\hbar}{2}$. Now, multiplying Eq. (2.17) on the right with $\langle\uparrow|$, we find:

$$\langle\uparrow|i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \langle\uparrow|\omega S_z|\psi(t)\rangle \implies i\hbar \frac{\partial}{\partial t} \langle\uparrow|\psi(t)\rangle = \omega\langle\uparrow|\frac{\hbar}{2}|\psi(t)\rangle \implies i\hbar \frac{\partial}{\partial t} \psi_\uparrow(t) = \frac{\hbar\omega}{2}\psi_\uparrow(t) \tag{2.18}$$

where $\psi_\uparrow(t) := \langle\uparrow|\psi(t)\rangle$ (the "spin-up" component of $|\psi(t)\rangle$). Analogously we find:

$$i\hbar \frac{\partial}{\partial t} \psi_\downarrow(t) = -\frac{\hbar\omega}{2}\psi_\downarrow(t) \tag{2.19}$$

In matrix form we can express this system of two ODEs (making the identification $|\uparrow\rangle \cong (1,0)^T, |\downarrow\rangle \cong (0,1)^T$) as:

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \psi_\uparrow(t) \\ \psi_\downarrow(t) \end{pmatrix} = \frac{\hbar\omega}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \psi_\uparrow(t) \\ \psi_\downarrow(t) \end{pmatrix} \tag{2.20}$$

These equations are easily solved by inspection to be complex exponentials (you can check!):

$$\psi_\uparrow(t) = c_\uparrow e^{-\frac{i\omega}{2}t}, \quad \psi_\downarrow(t) = c_\downarrow e^{\frac{i\omega}{2}t} \tag{2.21}$$

Where the coefficients $c_\uparrow, c_\downarrow$ are determined by the initial state $|\psi(0)\rangle$:

$$c_\uparrow = \langle\uparrow|\psi(0)\rangle = \frac{1}{\sqrt{2}} \tag{2.22}$$

$$c_\downarrow = \langle\downarrow|\psi(0)\rangle = \frac{1}{\sqrt{2}} \tag{2.23}$$

From which we conclude the solution to the SE in this case to be:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} e^{-\frac{i\omega}{2}t}|\uparrow\rangle + \frac{1}{\sqrt{2}} e^{\frac{i\omega}{2}t}|\downarrow\rangle. \tag{2.24}$$

Two comments before we move on; eigenstates of Hamiltonians do not evolve in time (except for a physically irrelevant phase factor $e^{-i\frac{E_n}{\hbar}t}$) - so if we instead started with state $|\psi(0)\rangle = |\uparrow\rangle$, the spin would be fully polarized upwards for all time and would not be influenced by the ($\hat{\mathbf{z}}$-oriented) external magnetic field. Yet again, we see that it is the relative phase that has physically observable effects.

On this note of eigenstates of a Hamiltonian begin stationary in time (hence sometimes called "stationary states"), we have the following generic recipe for solving the SE (in the case where we have a time-independent Hamiltonian H):

1. Solve for the eigenenergies E_n and corresponding eigenvectors $|n\rangle$ of H (Note - while this is generally fairly simple for small finite-dimensional systems, this is exceptionally difficult when we consider the continuous case (there are exceptionally few potentials that can be analytically solved), or systems with an infinite number of spins - for these we require approximation techniques, which we will discuss towards the end of the course). The evolution of these stationary states is just via the complex phase factor, e.g. $e^{-i\frac{E_n}{\hbar}t}|n\rangle$.

2. Expand out the initial state $|\psi(0)\rangle$ in the eigenbasis of H, i.e. $|\psi(0)\rangle = \sum_n \langle n|\psi(0)\rangle |n\rangle$ via resolution of the identity.

3. Solve for $|\psi(t)\rangle$ by considering the evolutions of each of the eigenstates independently, and then by linearity:

$$|\psi(t)\rangle = \sum_n \langle n|\psi(0)\rangle e^{-i\frac{E_n}{\hbar}t}|n\rangle. \tag{2.25}$$

Now, let's return back to our current example of the spin-1/2 particle in a magnetic field. Let's calculate the expectation value of S_x as a function of time:

$$\begin{aligned}
\langle S_x\rangle_\psi(t) &= \langle\psi(t)|S_x|\psi(t)\rangle \\
&= \left(\frac{1}{\sqrt{2}}e^{\frac{i\omega}{2}t}\langle\uparrow| + \frac{1}{\sqrt{2}}e^{-\frac{i\omega}{2}t}\langle\downarrow|\right) S_x \left(\frac{1}{\sqrt{2}}e^{-\frac{i\omega}{2}t}|\uparrow\rangle + \frac{1}{\sqrt{2}}e^{\frac{i\omega}{2}t}|\downarrow\rangle\right) \\
&= \left(\frac{1}{\sqrt{2}}e^{\frac{i\omega}{2}t}\langle\uparrow| + \frac{1}{\sqrt{2}}e^{-\frac{i\omega}{2}t}\langle\downarrow|\right) \frac{\hbar}{2} \left(\frac{1}{\sqrt{2}}e^{-\frac{i\omega}{2}t}|\downarrow\rangle + \frac{1}{\sqrt{2}}e^{\frac{i\omega}{2}t}|\uparrow\rangle\right) \\
&= \frac{\hbar}{4}\left(\langle\uparrow|\downarrow\rangle + e^{i\omega t}\langle\uparrow|\uparrow\rangle + e^{-i\omega t}\langle\downarrow|\downarrow\rangle + \langle\downarrow|\uparrow\rangle\right) \\
&= \frac{\hbar}{4}\left(e^{i\omega t} + e^{-i\omega t}\right) \\
&= \frac{\hbar}{2}\cos(\omega t)
\end{aligned} \tag{2.26}$$

where we have used that $S_x|\uparrow\ /\ \downarrow\rangle = \frac{\hbar}{2}|\downarrow\ /\ \uparrow\rangle$, the orthonormality of the up/down spin states and an Euler identity of $\frac{e^{i\omega t}+e^{-i\omega t}}{2} = \cos(\omega t)$. We see that the expectation value of S_x precesses in time! Note that the expectation value of S_y similarly precesses, with an analogous calculation showing that $\left\langle S_y\right\rangle_{\psi(t)} = \frac{\hbar}{2}\sin(\omega t)$. S_z commutes with the Hamiltonian and hence its expectation value is constant in time; and since $\langle\psi(0)|S_z|\psi(0)\rangle = \langle\rightarrow|S_z|\rightarrow\rangle = \frac{\hbar}{2}\langle\rightarrow|\leftarrow\rangle = 0$, $\langle S_z\rangle_{\psi(t)} = 0$ for all time.

We can visualize this precession of the spin using the *Bloch sphere* representation of a spin-1/2 (qubit) system. The state $|\psi(t)\rangle$ that we solved for in Eq. (2.24) can be visualized as a vector rotating along the surface of the unit sphere in the xy-plane:

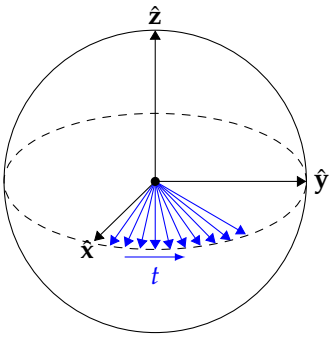

Figure 2.2: Bloch sphere visualization of spin precession; $|\psi(t)\rangle$ precesses in the xy-plane with frequency ω.

We will return to this Bloch sphere representation when we begin our discussion of quantum information; nevertheless, you might find the following exercise interesting; consider the spin operator in the

$\hat{\mathbf{n}}$ direction, given by $S_{\hat{\mathbf{n}}} = \mathbf{S} \cdot \hat{\mathbf{n}}$. Parameterizing in polar coordinates, we can write $\hat{\mathbf{n}} = (n_x, n_y, n_z)^T = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)^T$. You can check that the $(+\hbar/2)$ eigenvector for this operator can be solved to be:

$$|+\hat{\mathbf{n}}\rangle = \cos(\frac{\theta}{2})|\uparrow\rangle + e^{i\varphi}\sin(\frac{\theta}{2})|\downarrow\rangle \tag{2.27}$$

in the Bloch sphere representation, this can be realized as the vector with polar angle θ and azimuthal angle φ.

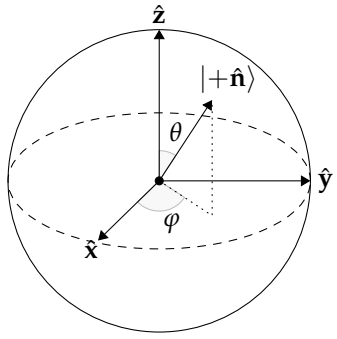

Figure 2.3: Bloch sphere representation of the spin-1/2 state $|+\hat{\mathbf{n}}\rangle = \cos(\frac{\theta}{2})|\uparrow\rangle + e^{i\varphi}\sin(\frac{\theta}{2})|\downarrow\rangle$, which lies on the surface of the unit two-sphere and has polar angle θ and azimuthal angle φ.

Note that Eq. (2.27) is a totally generic spin-1/2 state; naively, one needs two complex numbers, and hence four real parameters to specify the state of a spin-1/2 particle $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$. However, the normalization condition of $|\alpha|^2 + |\beta|^2 = 1$ and the irrelevancy of the global phase of $|\psi\rangle$ means that it is (physically) uniquely specified by two real numbers.

This example of spin precession is not only of pedagogical interest (being an example of where we can analytically solve the SE, and for arguably the simplest quantum system) but is also of practical interest - namely for medical imaging (i.e. NMR/MRI). When a (nuclear) spin precesses (under the influence of a **B**-field; which is why MRI requires such strong magnets) they emit electromagnetic radiation. This can then be picked up by an antenna, where the strength of the electromagnetic signal corresponds to the number of spins precessing, and therefore provides a measurement of the number of spins in a particular sample (e.g. of the brain).

We give an extremely high-level explanation here, as we have not yet developed all the tools necessary to discuss this in all the detail. We consider some sample of spins, where we want to measure the (spatially dependent) spin density. How would we go about doing this?

We begin with a static (time-independent) **B**-field, which we can WLOG take to be (as we have) in the $\hat{\mathbf{z}}$ direction, so $\mathbf{B} = B\hat{\mathbf{z}}$. Then the energy/Hamiltonian is given by:

$$E = -\frac{|e|B}{m_p c}S_z \tag{2.28}$$

From the above expression we can see that it is energetically favourable for the spins to align with the static **B**-field. Specifically, with $\omega = \frac{|e|B}{m_p c}$, we have that the spin-up states have eigenenergy $-\frac{\hbar\omega}{2}$, and the spin-down states have eigenenergy $\frac{\hbar\omega}{2}$. The two states have energy difference $\Delta E = \hbar\omega$.

The stronger this static magnetic field is, the more spins in our sample will be aligned with it (there will be some anti-aligned population, due to thermal excitations; the strength of the magnetic field then determines how much more likely the spins are to be in the ground (aligned) state, with the excited (anti-aligned) spins supressed by a Boltzmann factor $e^{-\hbar\omega/kT}$), and hence the stronger the net magnetization of

the spins. However, this static field alone cannot yield a precession signal, as the aggregate spin vector does not precess here; it is just aligned with the **B**-field. The individual spins may exhibit some precession about $\hat{\mathbf{z}}$ (as the quantum mechanical picture is that each spin in the sample is some superposition of the up/down spin states, and therefore can precess) but the phase of a given spin is random, and so when adding up to form the aggregate spin vector we find $\langle S_x \rangle = \left\langle S_y \right\rangle = 0$ and no precession is observed.

What we require then to obtain a precession signal is a second external magnetic field. Specifically, a field that is time-dependent, and rotating in the xy-direction with time:

$$\mathbf{B}(t) = B_1(\hat{\mathbf{x}}\cos(\omega' t) + \hat{\mathbf{y}}\sin(\omega' t)) \tag{2.29}$$

This time-dependent rotating field (which we do not know how to analyze yet - stay tuned for the treatment of time-dependent Hamiltonians and the Rabi formula, to come at the end of this course!) causes $p_\uparrow(t)$ and $p_\downarrow(t)$ to be time dependent (unlike the static $\mathbf{B} = B\hat{\mathbf{z}}$ field, which cannot cause changes in the up/down components of the spins in time); that is, we see spin "flips" in addition to spin precession from just the static magnetic field. Specifically, the probability of such spin flips is maximized when the rotating magnetic field frequency ω' matches the precession frequency ω, i.e. we have the resonance condition $\omega' = \omega$. You can picture this rotating magnetic field (when pulsed for an appropriate amount of time) as having the effect of "tipping" the aggregate magnetization vector into the xy-plane, and such the aggregate spin can precess and we are able to read out a (strong) EM signal from our sample.

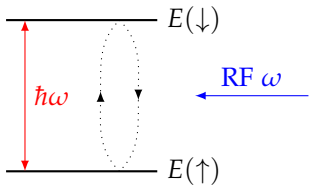

Figure 2.4: A RF (radiofrequency) field (rotating magnetic field) at the resonance frequency of $\omega' = \omega$ maximizes the probability of "spin flips" between the ground ($\uparrow$/ aligned) state to the excited ($\downarrow$/anti-aligned) state.

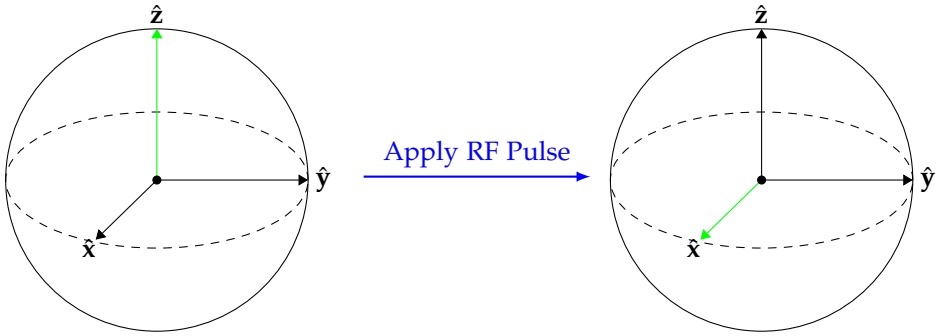

Figure 2.5: Applying a pulse of the (time-dependent) RF field causes the upwards-polarized (and therefore non-precessing) aggregate spin vector to tip into the xy-plane, where a maximal amount of precession signal can be read out.

How do we now introduce spatial resolution into this picture? Since the resonance frequency ω is dependent on the static field strength B, an approach is then to make the static B field position dependent;

this changes the resonance condition depending on where in the sample we are. This is illustrated in Fig. 2.6 below, where we have the static field strength varying in one dimension.

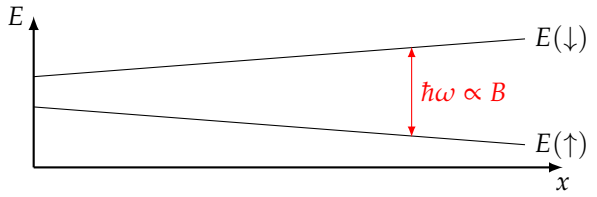

Figure 2.6: By applying a B-field with a gradient (here, increasing linearly as we go to the right) along the sample, the resonance frequency changes proportional to the B-field strength.

In three dimensions (of prime interest for imaging) we can use spatially dependent fields to get a signal from a specifically chosen 2D plane only. We can then vary this plane to get spatial resolution.

Let us go through an example in 2D; suppose we have a sample with spin-polarizable matter in the center. We can then consider varying the magnetic field over two planes, namely going vertically and horizontally through the sample. Based on the precession signal we extract (peaking at $x = y = 0$) we are able to resolve the sample.

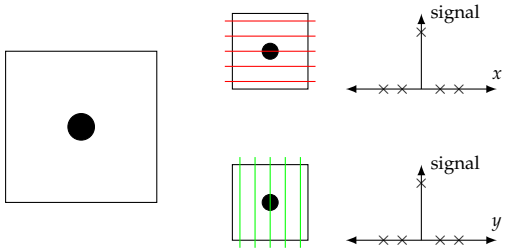

Figure 2.7: Simple example of 2D medical imaging. Suppose we have a sample with spin-polarizable matter in the center. By applying a vertical and horizontal gradient across the sample and measuring the signal strength, we are able to localize the spin-polarizable matter within the sample.

For general density patterns in 3D, we can choose multiple planes that go through the sample (over which we vary the B-field strength) in order to resolve it.

2.3 Unitarity of Schrödinger Evolution

We make an observation about our previous example of spin precession; we found that for intitial state $|\psi(0)\rangle = |+\rangle$ the state through time would be:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} e^{-\frac{i\omega}{2}t}|\uparrow\rangle + \frac{1}{\sqrt{2}} e^{\frac{i\omega}{2}t}|\downarrow\rangle. \tag{2.30}$$

and if we started with state $|\varphi(0)\rangle = |\uparrow\rangle$ the state through time would be:

$$|\varphi(t)\rangle = e^{-\frac{i\omega}{2}t}|\uparrow\rangle \tag{2.31}$$

and so we notice that:

$$\langle\varphi(t)|\psi(t)\rangle = \left(e^{\frac{i\omega}{2}t}\langle\uparrow|\right)\left(\frac{1}{\sqrt{2}} e^{-\frac{i\omega}{2}t}|\uparrow\rangle + \frac{1}{\sqrt{2}} e^{\frac{i\omega}{2}t}|\downarrow\rangle\right) = \frac{1}{\sqrt{2}}\langle\uparrow|\uparrow\rangle + \frac{1}{\sqrt{2}} e^{i\omega t}\langle\uparrow|\downarrow\rangle = \frac{1}{\sqrt{2}} \tag{2.32}$$

for all times t! That is, angles seem to be preserved under Schrödinger evolution. We can also visualize this angle preservation in the Bloch sphere representation; looking at Fig. 2.2, you can see that the angle (of $\theta = \pi/2$) between the $|\uparrow\rangle$ state (aligned along $\hat{\mathbf{z}}$) and the precessing spin in the xy-plane stays constant in time, even as the spin evolves.

Now, is this property generally true, or is this just a special case? It turns out that yes, it always holds for *any* two states $|\psi\rangle, |\varphi\rangle$ (which belong to an arbitrary Hilbert space) that Schrödinger evolution preserves the inner product:

$$\langle\varphi(t)|\psi(t)\rangle = \langle\varphi(0)|\psi(0)\rangle. \tag{2.33}$$

This statement is actually quite striking, given our experience in classical mechanics; given a classical ODE, two solutions can diverge in time (for example: consider the trajectory of an object thrown at two different initial velocities; even though the objects may start at the same position, the distance between them will grow apart in time). However, in QM the "distance" between states is fixed in time! A potential question that may arise: how do we explain chaos? Chaotic systems (that is, systems highly sensitive to initial conditions, where very slight differences can lead to vastly different future behaviour) are everywhere, for example the weather and N-body systems. Yet if quantum mechanics is the more correct theory underlying physics (where such diverging behaviour simply does not occur under evolution by the Schrödinger equation, as we will now prove), where does such chaotic behaviour arise?

We leave the reader to ponder the above question, and move onto the proof that Schrödinger evolution is unitary; from this the inner product preserving property will follow.

Theorem: Unitarity of Schrödinger evolution

Evolution under the Schrödinger equation is unitary. That is, given the time-evolution of a quantum state as goverened by the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H|\psi(t)\rangle \tag{2.34}$$

we can write $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ where $U(t)$ is the "time-evolution" operator. This time evolution operator is unitary.

Proof. First, note that writing $|\psi(t)\rangle = U(t)|\psi(0)\rangle$, we can substitute this into the Schrödinger equation to obtain:

$$i\hbar \frac{\partial}{\partial t} \left(U(t)|\psi(0)\rangle\right) = H(t)U(t)|\psi(0)\rangle \tag{2.35}$$

where we note that in general, the Hamiltonian may have time dependence. Since the above holds for arbitrary initial states $|\psi(0)\rangle$, it follows that:

$$i\hbar \frac{\partial}{\partial t} U(t) = H(t)U(t) \tag{2.36}$$

We now break up our analysis into cases.

1. H is constant in time. In this case, Eq. (2.36) is easily solved by inspection:

   $$U(t) = e^{-i\frac{H}{\hbar}t} \tag{2.37}$$

   But then taking the Hermitian conjugate, we have:

   $$U^\dagger(t) = e^{i\frac{H^\dagger}{\hbar}t} = e^{i\frac{H}{\hbar}t} \tag{2.38}$$

   as H is Hermitian. Therefore:

   $$U(t)U^\dagger(t) = U^\dagger(t)U(t) = \mathbb{I} \tag{2.39}$$

   and so U is unitary as claimed.

2. $\boxed{H \text{ depends on time}}$. For this case, note that (by rearranging Eq. (2.36)) that:

$$\frac{\partial}{\partial t} U(t) = \frac{H(t)}{i\hbar} U(t) \tag{2.40}$$

and taking the Hermitian conjugate of the above equation:

$$\frac{\partial}{\partial t} U^\dagger(t) = -\frac{U^\dagger(t)}{i\hbar} H(t) \tag{2.41}$$

where we use that $(AB)^\dagger = B^\dagger A^\dagger$ and that H is Hermitian. Now using the product rule:

$$\begin{aligned} \frac{\partial}{\partial t}\left(U^\dagger(t)U(t)\right) &= \left(\frac{\partial}{\partial t} U^\dagger(t)\right) U(t) + U^\dagger(t)\left(\frac{\partial}{\partial t} U(t)\right) \\ &= -\frac{1}{i\hbar} U^\dagger(t)H(t)U(t) + \frac{1}{i\hbar} U^\dagger(t)H(t)U(t) \\ &= 0 \end{aligned} \tag{2.42}$$

It therefore follows that $U^\dagger(t)U(t)$ is constant in time, and so $U^\dagger(t)U(t) = U^\dagger(0)U(0) = \mathbb{I}$ (as at time $t = 0$, the evolution operator does nothing!). The claim is therefore proven.

□

Corollary: Schrödinger evolution preserves the inner product

For any two quantum states $|\psi(t)\rangle, |\varphi(t)\rangle$ evolving under the Schrödinger equation:

$$\langle\varphi(t)|\psi(t)\rangle = \langle\varphi(0)|\psi(0)\rangle \tag{2.43}$$

Proof. Writing $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ and $|\varphi(t)\rangle = U(t)|\varphi(0)\rangle$, we have:

$$\langle\varphi(t)|\psi(t)\rangle = \langle\varphi(0)|U^\dagger(t)U(t)|\psi(0)\rangle = \langle\varphi(0)|\mathbb{I}|\psi(0)\rangle = \langle\varphi(0)|\psi(0)\rangle \tag{2.44}$$

where we have used unitarity in the second equality. □

2.4 The Heisenberg Picture and Ehrenfest's Theorem

So far, we have taken the view that quantum states $|\psi\rangle$ are the objects that evolve in time, and that observables are fixed through time. However, this is not the only way to do quantum mechanics. Often, we are interested in the expectation value of an observable A and how this evolves through time - Does it then make sense to consider observables as the objects that evolve through time, and the states as the fixed objects?

The answer turns out to be yes, and this is what is known as the Heisenberg picture. To see how things work in this picture, let's consider the expectation value of some observable A:

$$\begin{aligned} \langle A(t)\rangle_\psi &= \langle\psi(t)|A|\psi(t)\rangle \\ &= \left(\langle\psi(0)|U^\dagger(t)\right) A \left(U(t)|\psi(0)\rangle\right) \\ &= \langle\psi(0)|\left(U^\dagger(t)AU(t)\right)|\psi(0)\rangle \end{aligned} \tag{2.45}$$

We now define the observable A in the Heisenberg picture to be:

$$A^H(t) := U^\dagger(t)AU(t) \tag{2.46}$$

where A is the observable in the (familiar) Schrödinger picture, and $U(t)$ is the unitary time-evolution operator. And so, in the Heisenberg picture, the observable is what evolves through time!

If states evolve via the Schrödinger equation, what evolution equation does $A^H(t)$ satisfy? Let us now derive it.

Theorem: Heisenberg equation of motion

In the Heisenberg picture, the evolution of observables $A^H(t) := U^\dagger(t)AU(t)$ are governed by the Heisenberg equation of motion:

$$\frac{\mathrm{d}}{\mathrm{d}t}A^H(t) = \frac{1}{i\hbar}[A^H(t), H^H(t)] \tag{2.47}$$

where $H(t)$ is the Hamiltonian under which the system evolves (and $H^H = U^\dagger(t)H(t)U(t)$ its counterpart in the Heisenberg picture). In the case where H is independent of time (which is true for many cases of interest), $H^H(t) = H$ (i.e. the Hamiltonian is the same in the Schrödinger and Heisenberg pictures) and so the above simplifies to:

$$\frac{\mathrm{d}}{\mathrm{d}t}A^H(t) = \frac{1}{i\hbar}[A^H(t), H] \tag{2.48}$$

Proof. Using the definition of Heisenberg picture operators and the product rule, we have:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}A^H(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(U^\dagger(t)AU(t)\right)\\ &= \left(\frac{\mathrm{d}}{\mathrm{d}t}U^\dagger(t)\right)AU(t) + U^\dagger(t)A\left(\frac{\mathrm{d}}{\mathrm{d}t}U(t)\right)\end{aligned} \tag{2.49}$$

where we note that the derivative of the Schrödinger picture operators is zero since they are independent of time. Now, using Eqs. (2.40), (2.41) (derived from the Schrödinger equation) we find:

$$\frac{\mathrm{d}}{\mathrm{d}t}A^H(t) = -\frac{1}{i\hbar}U^\dagger(t)H(t)AU(t) + \frac{1}{i\hbar}U^\dagger(t)AH(t)U(t) \tag{2.50}$$

Inserting $\mathbb{I} = U(t)U^\dagger(t)$ in between the A and Hs:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}A^H(t) &= -\frac{1}{i\hbar}U^\dagger(t)H(t)U(t)U^\dagger(t)AU(t) + \frac{1}{i\hbar}U^\dagger(t)AU(t)U^\dagger(t)H(t)U(t)\\ &= -\frac{1}{i\hbar}\left(U^\dagger(t)H(t)U(t)\right)\left(U^\dagger(t)AU(t)\right) + \frac{1}{i\hbar}\left(U^\dagger(t)AU(t)\right)\left(U^\dagger(t)H(t)U(t)\right)\\ &= -\frac{1}{i\hbar}H^H(t)A^H(t) + \frac{1}{i\hbar}A^H(t)H^H(t)\\ &= \frac{1}{i\hbar}[A^H(t), H^H(t)]\end{aligned} \tag{2.51}$$

and so Eq. (2.47) has been shown. Now, in the case where H is time independent, we solved for $U(t)$ to be $U(t) = e^{-i\frac{H}{\hbar}t}$ (Eq. (2.37)), in which case it is clear to see that $U(t)$ commutes with H and so $H^H(t) = U^\dagger(t)HU(t) = U^\dagger(t)U(t)H = H$, and so the above reduces to:

$$\frac{\mathrm{d}}{\mathrm{d}t}A^H(t) = \frac{1}{i\hbar}[A^H(t), H] \tag{2.52}$$

which is precisely Eq. (2.48). □

If you've taken an advanced course in classical mechanics before, the Heisenberg equation of motion may look very familiar; there we had the classical equation of motion (for an observable $a(t)$):

$$\frac{\mathrm{d}}{\mathrm{d}t}\, a(t) = [a(t), H]_{\text{classical}} \tag{2.53}$$

where $[\cdot,\cdot]_{\text{classical}}$ is the Poisson bracket. Going from classical to quantum mechanics (the Heisenberg equation of motion), we promote observables to operator status, replace the Poisson bracket with a commutator, and introduce a factor of $i\hbar$. We saw this previously when we discussed the canonical commutation relations!

Let us explore this connection between classical mechanics and quantum mechanics further, by discussing Ehrenfest's theorem; before getting there however, there are some intermediate results we will require.

Lemma: Generalized position-momentum commutation relations

Let F, G be smooth functions. Then:

$$[X, F(P)] = i\hbar \frac{\partial F}{\partial p}(P) \tag{2.54}$$

$$[P, G(X)] = -i\hbar \frac{\partial G}{\partial x}(X) \tag{2.55}$$

Proof. Left as a homework exercise. However, let's check that the claim holds for the special cases of $F(P) = P$ and $F(P) = P^2$. If $F(P) = P$, then $[X, P] = i\hbar\mathbb{I}$ (canonical commutation relation) and $\frac{\partial F}{\partial p}(P) = \mathbb{I}$ so $i\hbar \frac{\partial F}{\partial p}(P) = i\hbar\mathbb{I}$ and so the claim holds. If $F(P) = P^2$, then we use the commutator identity:

$$[A, BC] = B[A, C] + [A, B]C \tag{2.56}$$

to obtain:

$$[X, P^2] = P[X, P] + [X, P]P = Pi\hbar\mathbb{I} + i\hbar\mathbb{I}P = 2i\hbar P \tag{2.57}$$

and since $\frac{\partial F}{\partial p}(P) = 2P$, then $i\hbar \frac{\partial F}{\partial p}(P) = 2i\hbar P$ and so the claim holds here as well. □

Before tackling Ehrenfest's Theorem in full generality, let us first consider the case of the free particle. Here, $V(X) = 0$ and so:

$$H = \frac{P^2}{2m} \tag{2.58}$$

Now, using the Heisenberg equation of motion, we find:

$$\frac{\mathrm{d}}{\mathrm{d}t} P = \frac{1}{i\hbar}[P, H] = 0 \implies P(t) = \text{constant}(t) = P(0) \tag{2.59}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} X = \frac{1}{i\hbar}[X, H] = \frac{i\hbar}{i\hbar}\frac{P}{M} = \frac{P}{M} \tag{2.60}$$

Where in the second equality we use that $[X, P^2] = 2i\hbar P$ as derived in Eq. (2.57) above. This is an easy first-order ODE with solution:

$$X(t) = X(0) + \frac{P(0)}{m} t \tag{2.61}$$

This is a very intriguing result; the equation of motion for the position of a classical free particle has the identical form of $x(t) = x_0 + \frac{p_0}{m}t$. Let's see how this generalizes when we introduce a potential V.

Theorem: Ehrenfest's Theorem

For a particle evolving under the Hamiltonian:

$$H = \frac{P^2}{2m} + V(X) \tag{2.62}$$

It's expectation value in position obeys:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle X\rangle = \frac{1}{m}\left\langle -\frac{\partial V}{\partial x}(X)\right\rangle. \tag{2.63}$$

Proof. Applying Heisenberg's equations of motion to calculate the time derivatives of momentum for the Hamiltonian given in Eq. (2.62), we have:

$$\frac{\mathrm{d}}{\mathrm{d}t}P = \frac{1}{i\hbar}[P,H] = \frac{1}{i\hbar}[P,V(X)] = -\frac{i\hbar}{i\hbar}\frac{\partial V}{\partial x}(X) = -\frac{\partial V}{\partial x}(X) \tag{2.64}$$

where we have used the generalized position-momentum commutation relations in the third equality. For position, we find the exact same result as in the free particle case:

$$\frac{\mathrm{d}}{\mathrm{d}t}X = \frac{1}{i\hbar}[X,H] = \frac{P}{m} \tag{2.65}$$

and so taking a time derivative of the above equation:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}X = \frac{1}{m}\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\partial V}{\partial x}(X) \tag{2.66}$$

where in the last equality we substitute Eq. (2.64). Now, taking expectation values, we obtain:

$$\left\langle \frac{\mathrm{d}^2}{\mathrm{d}t^2}X\right\rangle = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle X\rangle = \frac{1}{m}\left\langle -\frac{\partial V}{\partial x}(X)\right\rangle \tag{2.67}$$

which is the claimed result. □

Let's now interpret this formula we have derived; this looks very much like Newton's second law, with a second time derivative of $\langle X\rangle$ on the side, and $\left\langle -\frac{\partial V}{\partial x}(X)\right\rangle$ playing the role of the force. In fact in the limit of a delta peak (i.e. a "point particle") where $\psi(x) = \delta(x - x_0)$, we find that:

$$\left\langle -\frac{\partial}{\partial x}V(x)\right\rangle_\psi \to -\frac{\partial}{\partial x}V(x)\bigg|_{x_0} \tag{2.68}$$

and identifying $F(x_0) = -\frac{\partial V}{\partial x}(x_0)$, we see that Newton's law is reproduced exactly. However in general we observe that all $\hbar$s have vanished when we look at our derived formula, and $\langle X\rangle$ behaves classically (with the classical equations of motions precisely reproduced for narrow $\psi(x)$). Ehrenfest's theorem therefore gives us a concrete example of the quantum $\leftrightarrow$ classical correspondence.

2.5 The Quantum Harmonic Oscillator

Let's return back to the Schrödinger equation in the position basis, Eq. (2.3) (assuming that the potential is time-independent):

$$i\hbar \frac{\partial}{\partial t} \psi(x,t) = \left[V(x) - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right] \psi(x,t) \tag{2.69}$$

We now make a separation-of-variables ansatz, where:

$$\psi(x,t) = e^{-i\frac{E}{\hbar}t} \psi_E(x) \tag{2.70}$$

where $e^{-i\frac{E}{\hbar}t}$ is the time part of $\psi(x,t)$ and $\psi_E(x)$ is the spatial part. Substituting this into the position basis SE, we obtain the (familiar) time-independent Schrodinger equation (check!):

$$E\psi_E(x) = \left[V(x) - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right] \psi_E(x) \tag{2.71}$$

where E is the energy eigenvalue of ψ_E. Compare this to the basis-independent version, where $H|\psi\rangle = E|\psi\rangle$ for an energy eigenstate $|\psi\rangle$ of H. If we able to solve the time-independent SE for the eigenenergies and eigenstates $E/\psi_E(x)$, we have fully solved the time-evolution problem as we can express any initial state as the sum of energy eigenstates, which evolve via phase factor $e^{-i\frac{E}{\hbar}t}$ (recall the generic recipe for solving the SE which we discussed in Section 2.2):

$$\psi(x,t) = \sum_n \left(\int_{-\infty}^{\infty} \psi^*_{E_n}(x)\psi(0) \right) e^{-i\frac{E_n}{\hbar}t} \psi_n(x) \tag{2.72}$$

This gets us into the realm of solving bound state problems; no doubt you have encountered and solved many of these in a prior course, such as the square well potential, the hydrogen atom and the simple harmonic oscillator. It is worth noting that with the exception of these three (and the free particle and the delta function well) Eq. (2.71) cannot be solved analytically, and we have to resort to numerical techniques or approximation methods (stay tuned for the last unit of the course)!

Here, we will study the simple harmonic oscillator. Not only is it a practically relevant example as quadratic potentials come up everywhere in physics (for an arbitrary potential profile, most minima can be approximated to be quadratic), it is also a pedagogical primer, as in solving this problem we will use the algebraic technique of ladder operators which we will encounter again when we begin our study of angular momentum.

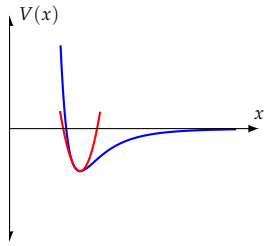

Figure 2.8: Plot of the Lennard Jones potential (a toy model for intermolecular interactions) $V(x) \sim \left(\left(\frac{1}{x}\right)^{12} - \left(\frac{1}{x}\right)^{6} \right)$ (blue) and approximation about the minima of the LJ potential by a quadratic potential $V(x) \sim x^2$ (red).

So, let's begin to analyze it! The Hamiltonian for the quantum harmonic oscillator is given by:

$$H = \frac{P^2}{2m} + \frac{1}{2} m\omega^2 X^2 \tag{2.73}$$

where $\omega = \sqrt{\frac{k}{m}}$ is the angular frequency of the analogous classical simple harmonic oscilaltor. We then define the annihilation and creation operators:

$$a \coloneqq \sqrt{\frac{m\omega}{2\hbar}} \left(X + i\frac{P}{m\omega} \right) \tag{2.74}$$

$$a^\dagger \coloneqq \sqrt{\frac{m\omega}{2\hbar}} \left(X - i\frac{P}{m\omega} \right) \tag{2.75}$$

the names given to these operators will soon become clear. Let's study their algebraic properties - first, let's calculate their commutator:

$$[a, a^\dagger] = \frac{m\omega}{2\hbar} \left(-i\frac{[X,P]}{m\omega} + i\frac{[P,X]}{m\omega} \right) = \mathbb{I} \tag{2.76}$$

where we have used the canonical commutation relation $[X, P] = i\hbar\mathbb{I}$. Further, note that:

$$a^\dagger a = \frac{m\omega}{2\hbar} \left(X^2 + \frac{P^2}{m^2\omega^2} \right) + \frac{i[X,P]}{2\hbar} = \frac{H}{\hbar\omega} - \frac{1}{2}\mathbb{I} \tag{2.77}$$

and so we can rewrite the Hamiltonian in terms of these operators as:

$$H = \hbar\omega(a^\dagger a + \frac{\mathbb{I}}{2}) \tag{2.78}$$

Now, let us define the number operator (again whose name will become clear in a moment):

$$N \coloneqq a^\dagger a \tag{2.79}$$

The commutator of this operator with the annihilation/creation operators is then:

$$[N, a] = [a^\dagger a, a] = a^\dagger \underbrace{[a, a]}_{0} + \underbrace{[a^\dagger, a]}_{-\mathbb{I}} a = -a \tag{2.80}$$

$$[N, a^\dagger] = a^\dagger \tag{2.81}$$

Given the definition of the number operator, we can write the QHO Hamiltonian as:

$$H = \hbar\omega(N + \frac{\mathbb{I}}{2}) \tag{2.82}$$

from which it is clear that H, N have a simultaneous eigenbasis. Let $\{|n\rangle\}$ be the eigenstates of N, with $N|n\rangle = n|n\rangle$. Then, using the commutation relations of N with a, we find:

$$\begin{aligned} N\left(a|n\rangle\right) &= (aN - a)|n\rangle \\ &= a(N - \mathbb{I})|n\rangle \\ &= (n-1)a|n\rangle \end{aligned} \tag{2.83}$$

In other words; if $|n\rangle$ is an eigenstate of N with eigenvalue n, then $a|n\rangle$ is an eigenstate of N with eigenvalue $n-1$ (up to some normalization factor):

$$a|n\rangle \propto |n-1\rangle \tag{2.84}$$

Analogously, we can show that:

$$a^\dagger|n\rangle \propto |n+1\rangle \tag{2.85}$$

from which we can see that if N counts the number of energy units of a state $|n\rangle$, a destroys one such energy unit (hence "annhilation operator") and $a^\dagger$ creates one.

We noted in our analysis above the presence of some normalization factor; let us solve for what this is. We want to find c in:

$$a|n\rangle = c|n-1\rangle \tag{2.86}$$

To this end, we consider taking an inner product of $a|n\rangle$ with itself:

$$\langle n|a^\dagger a|n\rangle = \langle n|N|n\rangle = n\langle n|n\rangle = n \tag{2.87}$$

but also:

$$\langle n|a^\dagger a|n\rangle = \langle n-1|c^* c|n-1\rangle = |c|^2\langle n-1|n-1\rangle = |c|^2 \tag{2.88}$$

From which we conclude comparing the two expressions that:

$$|c|^2 = n \tag{2.89}$$

This tells us that n must be real. By convention, we take c to be positive and real and so:

$$c = \sqrt{n} \tag{2.90}$$

Now, if we keep applying the annihilation operator to n, we obtain eigenkets of smaller and smaller eigenvalue:

$$\begin{aligned} a|n\rangle &= \sqrt{n}|n-1\rangle \\ a^2|n\rangle &= \sqrt{n(n-1)}|n-2\rangle \\ a^3|n\rangle &= \sqrt{n(n-1)(n-2)}|n-3\rangle \end{aligned} \tag{2.91}$$

so if we start with some non-negative integer n, then the sequence evenutally terminates as when we apply a n times, we end up with state $|0\rangle$, and if we apply it one more time, we find:

$$a^{n+1}|n\rangle = \sqrt{n(n-1)\dots(n-n)}|n-(n+1)\rangle = 0|-1\rangle = 0 \tag{2.92}$$

so the sequence terminates. If we started with some non-integer positive n, then this sequence no longer terminates; so it might seem like we could end up with eigenkets with $n < 0$. But this is not the case. Look back to Eq. (2.87); there we observed that:

$$\left(\langle n|a^\dagger\right)\left(a|n\rangle\right) = \langle n|N|n\rangle = n \tag{2.93}$$

but the inner product of any state of itself must be non-negative! From which we conclude that it is impossible for n to be negative; the sequence therefore *must* terminate with the $n = 0$ state, and therefore we conclude that n must be a non-negative integer.

So, we are done! The QHO has eigenstates $\left\{|n\rangle\right\}_{n=0}^{\infty}$, with associated eigenenergies:

$$H|n\rangle = \hbar\omega\left(N + \frac{\mathbb{I}}{2}\right)|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle \implies E_n = \hbar\omega\left(n + \frac{1}{2}\right) \tag{2.94}$$

There is a lot more that can be (easily) done here using this algebraic approach to the QHO. Let's look at a couple. First, note that analogously to finding $a|n\rangle = \sqrt{n}|n-1\rangle$, we find:

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{2.95}$$

Since the eigenstates $|n\rangle$ are orthonormal, we have the relations:

$$\langle n'|a|n\rangle = \sqrt{n}\delta_{n',n-1}, \quad \langle n'|a^\dagger|n\rangle = \sqrt{n+1}\delta_{n',n+1} \tag{2.96}$$

And if we rewrite X, P in terms of the ladder operators (inverting Eqs. (2.74), (2.75)):

$$X = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger) \tag{2.97}$$

$$P = i\sqrt{\frac{m\hbar\omega}{2}}(-a + a^\dagger) \tag{2.98}$$

we can easily solve for the matrix elements of position/momentum to be:

$$\langle n'|X|n\rangle = \sqrt{\frac{\hbar}{2m\omega}}(\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}) \tag{2.99}$$

$$\langle n'|P|n\rangle = i\sqrt{\frac{m\hbar\omega}{2}}(-\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}) \tag{2.100}$$

Taking $n' = n$ in Eq. (2.99), we find for any n that:

$$\langle X\rangle_n = \langle n|X|n\rangle = 0 \tag{2.101}$$

So all eigenstates of the QHO have expectation value zero for position[16]! The wavefunctions are either symmetric or anti-symmetric about zero. Identically:

$$\langle P\rangle_n = 0. \tag{2.102}$$

Now, to find the expectation value of X^2 we can calculate:

$$\begin{aligned}
\left\langle X^2\right\rangle_n &= \langle n|X^2|n\rangle \\
&= \frac{\hbar}{2m\omega}\langle n|a^2 + (a^\dagger)^2 + aa^\dagger + a^\dagger a|n\rangle \\
&= \frac{\hbar}{2m\omega}\langle n|a^2 + (a^\dagger)^2 + \underbrace{a^\dagger a}_{N} + \mathbb{I} + \underbrace{a^\dagger a}_{N}|n\rangle \\
&= \frac{\hbar}{2m\omega}\left(\sqrt{n(n-1)}\langle n|n-2\rangle + \sqrt{(n+1)(n+2)}\langle n|n+2\rangle + 2n\langle n|n\rangle + \langle n|n\rangle\right) \\
&= \frac{\hbar}{m\omega}(n + \frac{1}{2})
\end{aligned} \tag{2.103}$$

where we use $[a, a^\dagger] = \mathbb{I}$ in the third equality. Analogously:

$$\left\langle P^2\right\rangle_n = m\hbar\omega(n + \frac{1}{2})^2 \tag{2.104}$$

So the products of the variances for the eigenstates are:

$$\left\langle (\Delta X)^2\right\rangle_n \left\langle (\Delta P)^2\right\rangle_n = (\left\langle X^2\right\rangle_n - \langle X\rangle_n^2)(\left\langle P^2\right\rangle_n - \langle P\rangle_n^2) = \hbar^2\left(n + \frac{1}{2}\right)^2 \tag{2.105}$$

and we note that when $n = 0$ (the ground state) that:

$$\left\langle (\Delta X)^2\right\rangle_0 \left\langle (\Delta P)^2\right\rangle_0 = \frac{\hbar^2}{4} \tag{2.106}$$

[16] We will give an alternative proof of this fact using parity operators when we discuss symmetry later in the course.

i.e. the Heisenberg uncertainty principle is saturated! This tells us that the ground state position wavefunction is a Gaussian (as we showed in Chapter 1). Let us verify this statement using a different approach; the ground state of the QHO is defined by the equation:

$$a|0\rangle = 0 \tag{2.107}$$

Expanding this out in the position basis, we have:

$$\langle x|a|0\rangle = \langle x|\sqrt{\frac{m\omega}{2\hbar}}\left(X + i\frac{P}{m\omega}\right)|0\rangle = \sqrt{\frac{m\omega}{2\hbar}}\left(x\langle x|0\rangle + \frac{i}{m\omega}\left(-i\hbar\frac{\partial}{\partial x}\right)\langle x|0\rangle\right) = 0 \tag{2.108}$$

where we have used the representation of momentum in the position basis (Eq. (1.117)). This gives us a differential equation for $\psi_0(x) = \langle x|0\rangle$:

$$x\psi_0(x) + \frac{\hbar}{m\omega}\frac{\partial}{\partial x}\psi_0(x) = 0 \tag{2.109}$$

which is solved by:

$$\psi_0(x) = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4}\exp\left(-\frac{1}{2}\left(\frac{m\omega x^2}{\hbar}\right)\right) \tag{2.110}$$

which is a Gaussian, as we expected! The excited state wavefunctions can be derived by taking the excited state kets:

$$|n\rangle = \left[\frac{(a^\dagger)^n}{\sqrt{n!}}\right]|0\rangle \tag{2.111}$$

and projecting them onto the position basis.

3 Quantum Information and Foundations

3.1 Composite Systems and Entanglement

We so far have discussed the quantum mechanics of a single particle, where the states of the particle live in a Hilbert space $\mathcal{H}$. How then do we discuss the quantum mechanics of multiple particles? We introduce the tensor product, which is the operation that allows us to compose systems in quantum mechanics[17].

Definition: Tensor Product of States/Operators

Given states $|\alpha\rangle_A \in \mathcal{H}_A$ and $|\beta\rangle_B \in \mathcal{H}_B$, we can compose them using the tensor product:

$$|\alpha\rangle_A \otimes |\beta\rangle_B \tag{3.1}$$

The tensor product has the distributive property:

$$|\alpha\rangle_A \otimes (|\beta_1\rangle_B + |\beta_2\rangle_B) = |\alpha\rangle_A \otimes |\beta_1\rangle_B + |\alpha\rangle_A \otimes |\beta_2\rangle_B \tag{3.2}$$

It also has the property that the inner product of two tensor product states is just the product of the inner products on each of the subspaces:

$$\left(\langle\alpha|_A \otimes \langle\beta|_B\right)\left(|\gamma\rangle_A \otimes |\delta\rangle_B\right) = \langle\alpha|\gamma\rangle\langle\beta|\delta\rangle \tag{3.3}$$

Furthermore, the action of a composite linear operator $O_A \otimes P_B$ (where O_A acts on states in $\mathcal{H}_A$ and P_B acts on states in $\mathcal{H}_B$) on a composite state $|\alpha\rangle_A \otimes |\beta\rangle_B$ is to act on each of the states in the respective subspaces, and then take the tensor product:

$$(O_A \otimes P_B)(|\alpha\rangle_A \otimes |\beta\rangle_B) = (O_A|\alpha\rangle_A) \otimes (P_B|\beta\rangle_B) \tag{3.4}$$

Note that the tensor product is not commutative; $|\alpha\rangle_A \otimes |\beta\rangle_B \neq |\beta\rangle_B \otimes |\alpha\rangle_A$ in general.

It may help to consider what the tensor product does on the level of vector/matrix representations of states/operators (in the finite-dimensional case). Suppose that we have the following representations of $|\alpha\rangle_A, |\beta\rangle_B, O_A, P_B$:

$$|\alpha\rangle_A \cong [\alpha_A] = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}, \quad |\beta\rangle_B \cong [\beta_B] = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \tag{3.5}$$

$$O_A \cong [O_A] = \begin{pmatrix} O_{11} & O_{12} & \cdots & O_{1m} \\ O_{21} & O_{22} & & \\ \vdots & & \ddots & \\ O_{m1} & & & O_{mm} \end{pmatrix}, \quad P_B \cong [P_B] = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1k} \\ P_{21} & P_{22} & & \\ \vdots & & \ddots & \\ P_{k1} & & & P_{kk} \end{pmatrix} \tag{3.6}$$

Then the representations of $|\alpha\rangle_A \otimes |\beta\rangle_B$, $O_A \otimes P_B$ take the form:

$$|\alpha\rangle_A \otimes |\beta\rangle_B \cong \begin{pmatrix} \alpha_1[\beta] \\ \alpha_2[\beta] \\ \vdots \\ \alpha_m[\beta] \end{pmatrix} \tag{3.7}$$

[17] It is worth noting that even to describe the quantum mechanics of a single particle, we require this composition operation - the Hilbert spaces corresponding to the position and spin degrees of freedom of a single particle compose via the tensor product to yield a complete description of the particle.

$$O_A \otimes P_B \cong \begin{pmatrix} O_{11}[P_B] & O_{12}[P_B] & \cdots & O_{1m}[P_B] \\ O_{21}[P_B] & O_{22}[P_B] & & \\ \vdots & & \ddots & \\ O_{m1}[P_B] & & & O_{mm}[P_B] \end{pmatrix} \tag{3.8}$$

From this we get the intuition for the fact that if $|\alpha\rangle_A$ is an m-dimensional state ket and $|\beta\rangle_B$ is an k-dimensional state ket, then $|\alpha\rangle_A \otimes |\beta\rangle_B$ is mk-dimensional.

Of course, the composite states $|\alpha\rangle_A \otimes |\beta\rangle_B$ are themselves quantum states, so they must live in some Hilbert space ; this larger/composite Hilbert space (which the above analysis suggests should have dimension given by $\dim(\mathcal{H}_A) \cdot \dim(\mathcal{H}_B)$) is constructed as follows:

Definition: Tensor Product Hilbert Space

Let $\{|i\rangle_A\}_{i=1}^m, \{|j\rangle_B\}_{j=1}^k$ be ONBs for Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$ (with $\dim(\mathcal{H}_A) = m$ and $\dim(\mathcal{H}_B) = k$). The composite Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ is then defined as the span of the basis states:

$$|ij\rangle = |i\rangle_A \otimes |j\rangle_B \tag{3.9}$$

That is to say; the states Hilbert space $\mathcal{H}_{AB}$ consists of all (normalized) complex linear combinations $|\psi\rangle_{AB}$:

$$|\psi\rangle_{AB} = \sum_{ij} c_{ij}|ij\rangle, \quad c_{ij} \in \mathbb{C}. \tag{3.10}$$

Note that although we have used a particular basis to define it, $\mathcal{H}_{AB}$ is a basis-independent construction (check!)

Although we have here focused on the case of composing two systems, the above definitions generalize to the case when we compose n quantum mechanical systems together - for example we may have n particles whose states live in Hilbert spaces $\mathcal{H}_i$ and the n-particle Hilbert space is given by:

$$\mathcal{H}_{\text{composite}} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n \tag{3.11}$$

which is defined as the span of the vectors:

$$|j^{(1)}\rangle_1 \otimes |j^{(2)}\rangle_2 \otimes \ldots \otimes |j^{(n)}\rangle_n \tag{3.12}$$

where each $|j^{(i)}\rangle_i$ belongs to $\left\{|j^{(i)}\rangle_i\right\}_{j=1}^{\dim(\mathcal{H}_i)}$, an ONB of $\mathcal{H}_i$.

Note that the tensor product is associative; that is:

$$\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C = \mathcal{H}_A \otimes (\mathcal{H}_B \otimes \mathcal{H}_C) = (\mathcal{H}_A \otimes \mathcal{H}_B) \otimes \mathcal{H}_C. \tag{3.13}$$

An analogous associative property holds for the tensor product of states and operators - so, the n-fold tensor product can just be viewed as just iterating the tensor product for the case of 2 systems n times.

Further note - the number of basis vectors of the composite Hilbert space (and hence its dimension) is given by $\dim(\mathcal{H}_{\text{composite}}) = \prod_{i=1}^n \dim(\mathcal{H}_i)$ - this is exponential in the number of systems being composed. For example for n spin-1/2 particles (with $\dim(\mathcal{H}_i) = 2$), the dimension of the composite Hilbert space is $\dim(\mathcal{H}^n) = \prod_{i=1}^n 2 = 2^n$. For $n = 300$ we have $\dim(\mathcal{H}^n) \sim 10^{90}$ which already exceeds the number of atoms in the observable universe ($10^{78} - 10^{82}$). This high dimensionality is a reason[18] for why quantum systems are hard to simulate classically.

[18]There are some subtleties here; specifically, we require an extremely large number of parameters to describe highly entangled states (entanglement to be defined extremely shortly). Product (i.e. unentangled) states, i.e. states of the form in Eq. (3.12) are efficiently simulable because we may describe the subsystems individually, and therefore the whole state efficiently. The argument is actually a layer more nuanced than this, because certain types of entangled states (stabilizer states - see the Gottesman-Knill Theorem) are efficiently simulable. But this is far beyond the scope of this course.

Let us give a concrete example of $n = 2$ spin-1/2 particles. An ONB for the Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$ is $\{|\uparrow\rangle, |\downarrow\rangle\}$, so the basis states of the composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ are:

$$\begin{aligned} |\uparrow\uparrow\rangle_{AB} &:= |\uparrow\rangle_A \otimes |\uparrow\rangle_B \\ |\uparrow\downarrow\rangle_{AB} &:= |\uparrow\rangle_A \otimes |\downarrow\rangle_B \\ |\downarrow\uparrow\rangle_{AB} &:= |\downarrow\rangle_A \otimes |\uparrow\rangle_B \\ |\downarrow\downarrow\rangle_{AB} &:= |\downarrow\rangle_A \otimes |\downarrow\rangle_B \end{aligned} \tag{3.14}$$

And so:

$$\mathcal{H}_{AB} = \text{span}(\{|\uparrow\uparrow\rangle_{AB}, |\uparrow\downarrow\rangle_{AB}, |\downarrow\uparrow\rangle_{AB}, |\downarrow\downarrow\rangle_{AB}\}) = \{\alpha|\uparrow\uparrow\rangle_{AB} + \beta|\uparrow\downarrow\rangle_{AB} + \gamma|\downarrow\uparrow\rangle_{AB} + \delta|\downarrow\downarrow\rangle_{AB} : \alpha, \beta, \gamma, \delta \in \mathbb{C}\} \tag{3.15}$$

A question we now ask - are all states in a composite Hilbert space able to be written as a tensor product of states of the individual subsystems (as the notation $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ might suggest)? The answer is a *no* - this leads to our definition of entanglement, which will play a key role in the entire discussion of this chapter:

Definition: Entanglemement

A pure quantum state $|\Psi\rangle$ in a composite Hilbert space $\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i$ is *entangled* if it cannot be written as the tensor product of states from the subsystems $\mathcal{H}_1, \ldots \mathcal{H}_n$, i.e.:

$$|\Psi\rangle \neq |\psi_1\rangle_1 \otimes |\psi_2\rangle_2 \otimes \ldots \otimes |\psi_n\rangle_n \tag{3.16}$$

for *any* choice of states $|\psi_i\rangle_i \in \mathcal{H}_i$.

For the case of $n = 2$ subsystems, we have bipartite entanglement defined as:

Definition: Bipartite entanglement

Let $\mathcal{H}_A, \mathcal{H}_B$ be Hilbert spaces and define the composite Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. A pure state $|\Psi\rangle_{AB} \in \mathcal{H}_{AB}$ is *entangled* if:

$$|\Psi\rangle_{AB} \neq |\psi\rangle_A \otimes |\varphi\rangle_B \tag{3.17}$$

for any choice of local states $|\psi\rangle_A \in \mathcal{H}_A, |\varphi\rangle_B \in \mathcal{H}_B$.

A specific example of bipartite entanglement is given by the Bell state $|B_{11}\rangle$ (also called the singlet state - this name for it will perhaps become clearer after we begin our study of addition of angular momenta):

$$|B_{11}\rangle = \frac{|\uparrow\rangle_A \otimes |\downarrow\rangle_B - |\downarrow\rangle_A \otimes |\uparrow\rangle_B}{\sqrt{2}} \tag{3.18}$$

It is a useful exercise to use the definition of entanglement given above to prove that the above Bell state is indeed entangled (hint: try a proof by contradiction).

Let's explore some properties of this state - let us begin by looking at what happens when we measure one of the two spins. In general, if we perform an operation on one subsystem (represented by the application of an operator A) of a composite system while doing nothing to the other parts, we can represent this by the composite operator consisting of applying A to the specific subsystem, tensored with the identity operation $\mathbb{I}$ on the other subsystems. In our case, we consider operators of the form $\Pi_A \otimes \mathbb{I}_B$ where Π_A is a projector acting on the first spin.

Let's suppose we measure the first spin in the $\{|\uparrow\rangle_A, |\downarrow\rangle_A\}$ basis. From the Born rule we find:

$$\begin{aligned} p(\uparrow) &= \langle B_{11}|\Pi_{\uparrow,A} \otimes \mathbb{I}_B|B_{11}\rangle \\ &= \frac{\langle\uparrow|\Pi_\uparrow|\uparrow\rangle\langle\downarrow|\mathbb{I}|\downarrow\rangle + \langle\downarrow|\Pi_\uparrow|\downarrow\rangle\langle\uparrow|\mathbb{I}|\uparrow\rangle - \langle\uparrow|\Pi_\uparrow|\downarrow\rangle\langle\downarrow|\mathbb{I}|\uparrow\rangle - \langle\downarrow|\Pi_\uparrow|\uparrow\rangle\langle\uparrow|\mathbb{I}|\downarrow\rangle}{2} \\ &= \frac{1\cdot 1 + 0\cdot 1 - 0\cdot 0 - 0\cdot 0}{2} \\ &= \frac{1}{2} \end{aligned} \tag{3.19}$$

and analogously $p(\downarrow) = \frac{1}{2}$. The Dirac postulate tells us that if we measure spin-up, then the post-measurement state is:

$$|B_{11}\rangle \to \frac{\Pi_{\uparrow,A} \otimes \mathbb{I}_B|B_{11}\rangle}{\sqrt{\langle B_{11}|\Pi_{\uparrow,A} \otimes \mathbb{I}_B|B_{11}\rangle}} = \frac{1}{\sqrt{\frac{1}{2}}} \frac{\Pi_{\uparrow,A}|\uparrow\rangle_A \otimes \mathbb{I}_B|\downarrow\rangle_B - \Pi_{\uparrow,A}|\downarrow\rangle_A \otimes \mathbb{I}_B|\uparrow\rangle_B}{\sqrt{2}} = |\uparrow\rangle_A \otimes |\downarrow\rangle_B \tag{3.20}$$

Analogously, it can be shown that if we measure the first spin to be spin-down, then the post-measurement state is:

$$|B_{11}\rangle \to |\downarrow\rangle_A \otimes |\uparrow\rangle_B \tag{3.21}$$

We note two things - it seems as though when we measure the first spin in the S_z eigenbasis that we have a 50/50 probability of measuring the first spin to be up or down, and that the second spin after the measurement points in the direction opposite that of which the first spin was measured to be.

Perhaps this interesting result is just a consequence of our choice of measurement basis $\{|\uparrow\rangle, |\downarrow\rangle\}$. Let us try then measuring in the S_x eigenbasis of $\left\{|\rightarrow\rangle = \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}, |\leftarrow\rangle = \frac{|\uparrow\rangle-|\downarrow\rangle}{\sqrt{2}}\right\}$. Using that $|\uparrow \ / \ \downarrow\rangle = \frac{|\rightarrow\rangle \pm |\leftarrow\rangle}{\sqrt{2}}$, we can rewrite the $|B_{11}\rangle$ state in terms of the S_x eigenstates as:

$$\begin{aligned} |B_{11}\rangle &= \frac{\frac{|\rightarrow\rangle_A+|\leftarrow\rangle_A}{\sqrt{2}} \otimes \frac{|\rightarrow\rangle_B-|\leftarrow\rangle_B}{\sqrt{2}} - \frac{|\rightarrow\rangle_A-|\leftarrow\rangle_A}{\sqrt{2}} \otimes \frac{|\rightarrow\rangle_B+|\leftarrow\rangle_B}{\sqrt{2}}}{\sqrt{2}} \\ &= \frac{|\leftarrow\rangle_A \otimes |\rightarrow\rangle_B - |\rightarrow\rangle_A \otimes |\leftarrow\rangle_B}{\sqrt{2}} \end{aligned} \tag{3.22}$$

Up to a (physically irrelevant) global minus sign, the form of $|B_{11}\rangle$ expressed in terms of S_x eigenstates is identical to $|B_{11}\rangle$ expressed in terms of S_z eigenstates. So, if we were to measure the first spin in the S_x eigenbasis, just as before, we would find that we would have 50/50 probability of measuring spin right or spin left, and the post-measurement state would have the unmeasured spin pointing in the opposite direction as the measured one.

In fact we could go through with the above calculation for an arbitrary measurement basis, and find the same result.

Proposition: $|B_{11}\rangle$ is non-local and anti-correlated in every direction

Consider the Bell state $|B_{11}\rangle$:

$$|B_{11}\rangle = \frac{|\uparrow\rangle_A \otimes |\downarrow\rangle_B - |\downarrow\rangle_A \otimes |\uparrow\rangle_B}{\sqrt{2}} \tag{3.23}$$

and consider an arbitrary ONB (for a spin-1/2 particle):

$$\mathcal{B}(\alpha, \beta) = \left\{ |\mathbf{r}_{\alpha,\beta}\rangle := \alpha|\uparrow\rangle + \beta|\downarrow\rangle, |\bar{\mathbf{r}}_{\alpha.\beta}\rangle := \beta^*|\uparrow\rangle - \alpha^*|\downarrow\rangle \right\} \tag{3.24}$$

where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2$. Then:

1. $|B_{11}\rangle$ has no local properties; that is, whatever parameters $\alpha, \beta \in \mathbb{C}$ the measurement of particle A in the basis $\mathcal{B}(\alpha, \beta)$ leads to a 50/50 distribution of outcome.
2. The measurement of particle A leads to the post-measurement states:

$$\begin{aligned} \text{outcome "+": } |B_{11}\rangle &\to |\mathbf{r}_{\alpha,\beta}\rangle_A \otimes |\bar{\mathbf{r}}_{\alpha,\beta}\rangle_B \\ \text{outcome "-": } |B_{11}\rangle &\to |\bar{\mathbf{r}}_{\alpha,\beta}\rangle_A \otimes |\mathbf{r}_{\alpha,\beta}\rangle_B \end{aligned} \tag{3.25}$$

   That is, after measurement, the spin states of particles A/B are perfectly anti-correlated, irrespective of the measurement outcome and measurement basis.

The above demonstrates how quantum entanglement can give rise to "stronger-than-classical" correlations. In classical mechanics, it is possible for measurements to be correlated in certain measurement bases, but not all.

Proof. Left as a homework exercise. □

3.2 The No-Cloning Theorem

Consider now a thought experiment where, upon preparing two spin-1/2 particles in a Bell state $|B_{11}\rangle$, we flew out one pair to the moon while we kept one on Earth. Suppose we were to measure the particle on Earth; then the formalism of quantum mechanics would tell us that the particle on the moon would instantaneously collapse to the spin state pointing opposite to that which was measured on Earth. This phenomenon, coined "spooky action at a distance" by Einstein, seems quite troubling; it seems as information is travelling faster than the speed of light when the measurement is made! Could we harness this quantum-mechanical effect to communicate superliminally (and therefore - break the laws of special relativity)?

There is no need to fear, as the answer is no, as we will prove this in full generality in the latter half of this chapter. However, it will be of interest to consider a specific example of a protocol which does not work, as the reason for its failure is highly interesting. The (non) protocol for superluminal communication goes as follows:

(Non)-Protocol: Superluminal communication

Objective: Transmit one bit b of information (where $b \in \mathbb{Z}_2 = \{0,1\}$) from A to B.
Protocol:

1. In advance of the actual communication, the parties share a Bell state $|B_{11}\rangle$ between themselves.
2. If $b = 0$, then A measures their particle of $|B_{11}\rangle$ in the eigenbasis of S_z, i.e. in $\mathcal{B}_z = \{|\uparrow\rangle, |\downarrow\rangle\}$. If $b = 1$, then A measures their particle of $|B_{11}\rangle$ in the eigenbasis of S_x, i.e. in $\mathcal{B}_x = \{|\rightarrow\rangle, |\leftarrow\rangle\}$.
3. Party B copies their state a large number of times, i.e.:

$$\begin{aligned} |\uparrow\rangle &\rightarrow |\uparrow\rangle \otimes |\uparrow\rangle \otimes |\uparrow\rangle \otimes \dots \\ |\downarrow\rangle &\rightarrow |\downarrow\rangle \otimes |\downarrow\rangle \otimes |\downarrow\rangle \otimes \dots \\ |\rightarrow\rangle &\rightarrow |\rightarrow\rangle \otimes |\rightarrow\rangle \otimes |\rightarrow\rangle \otimes \dots \\ |\leftarrow\rangle &\rightarrow |\leftarrow\rangle \otimes |\leftarrow\rangle \otimes |\leftarrow\rangle \otimes \dots \end{aligned} \tag{3.26}$$

4. B identifies the state received in the transmission, through the measurement of the multiple copies (doing a large number of measurements in the S_z and S_x eigenbases, until they are confident that the state transmitted is one of $|\uparrow\rangle/|\downarrow\rangle$ or $|\rightarrow\rangle/|\leftarrow\rangle$). If the state received was $|\uparrow\rangle$ or $|\downarrow\rangle$, B outputs $b = 0$. If the state received was $|\leftarrow\rangle$ or $|\rightarrow\rangle$, B outputs $b = 1$.

Where does this protocol fail? The first and second steps are fine; there is no problem with creating a Bell state, then taking them far apart from each other, and measuring one of the two spins in a particular basis. The fourth step is also fine; if we are given a large number of identical states, by doing a sufficiently large number of measurements (in different bases), we can be confident about the state that we have (and hence obtain the correct output). The failure of the protocol comes in step 3 - namely, *unknown quantum states cannot be copied.* This is known as the *No-cloning theorem*, which is simple to prove but has profound implications.

Theorem: No-cloning

Let $|\psi\rangle \in \mathcal{H}_d$ be an unknown quantum state, and let $|0\rangle \in \mathcal{H}_d$ be a fixed known quantum state[a]. Then, the copying (cloning) operation C defined by:

$$C : |\psi\rangle \otimes |0\rangle \rightarrow |\psi\rangle \otimes |\psi\rangle, \quad \forall |\psi\rangle \in \mathcal{H}_d \tag{3.27}$$

cannot be realized in quatnum mechanics.

[a] Note: $|0\rangle$ is *not* the null ket

The proof of the above theorem rests on the linearity of quantum mechanics.

Lemma: Linearity of quantum mechanics

Unitary evolution according to the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H|\psi(t)\rangle \tag{3.28}$$

is linear; that is, if $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$ solve Eq. (3.28), then:

$$|\psi_{a,b}(t)\rangle = a|\psi_1(t)\rangle + b|\psi_2(t)\rangle \tag{3.29}$$

is also a solution to Eq. (3.28). In addition, measurement according to the Dirac projection postulate is linear up to normalization:

$$\Pi_i(a|\psi_1(t)\rangle + b|\psi_2(t)\rangle) = a\Pi_i|\psi_1(t)\rangle + b\Pi_i|\psi_2(t)\rangle. \tag{3.30}$$

Proof. Plugging in $|\psi_{a,b}(t)\rangle$ in to the LHS of Eq. (3.28), we find:

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} |\psi_{a,b}(t)\rangle &= i\hbar \frac{\partial}{\partial t} (a|\psi_1(t)\rangle + b|\psi_2(t)\rangle) \\ &= ai\hbar \frac{\partial}{\partial t} |\psi_1(t)\rangle + bi\hbar \frac{\partial}{\partial t} |\psi_2(t)\rangle \\ &= aH|\psi_1(t)\rangle + bH|\psi_2(t)\rangle \\ &= H(a|\psi_1(t)\rangle + b|\psi_2(t)\rangle) \\ &= H|\psi_{a,b}(t)\rangle \end{aligned} \tag{3.31}$$

where in the second equality we use the linearity of the derivative, in the third equality we use that $|\psi_1(t)\rangle, |\psi_2(t)\rangle$ are individually solutions to Eq. (3.28), and in the fourth equality we use the linearity of the Hamiltonian operator H. We have thus shown that $|\psi_{a,b}(t)\rangle$ is also a solution to Eq. (3.28).

Next, the Dirac projection postulate tells us that:

$$|\psi\rangle \to \frac{\Pi_i|\psi\rangle}{\sqrt{\langle\psi|\Pi_i|\psi\rangle}} \tag{3.32}$$

So neglecting the normalization factor:

$$|\psi\rangle \to \propto \Pi_i|\psi\rangle \tag{3.33}$$

Therefore since projectors are linear:

$$\begin{aligned} |\psi_{a,b}\rangle \to &\propto \Pi_i|\psi_{a,b}\rangle \\ &= a\Pi_i|\psi_1\rangle + b\Pi_i|\psi_2\rangle \end{aligned} \tag{3.34}$$

which proves the claim. □

Having made this observation about linearity, we can proceed to the proof of the no-cloning theorem.

Proof. Assume for the sake of contradiction that C exists. We have $|0\rangle \in \mathcal{H}_d$ the fixed/reference quantum state, and let $|1\rangle \in \mathcal{H}_d$ be some state orthogonal to $|0\rangle$. By assumption, C clones $|0\rangle$ and $|1\rangle$, so:

$$\begin{aligned} C(|0\rangle \otimes |0\rangle) &\propto |0\rangle \otimes |0\rangle \\ C(|1\rangle \otimes |0\rangle) &\propto |1\rangle \otimes |1\rangle \end{aligned} \tag{3.35}$$

furthermore, defining $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$, C should also clone this state:

$$C(|+\rangle \otimes |0\rangle) \propto |+\rangle \otimes |+\rangle = \frac{|0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle}{2} \tag{3.36}$$

However, since C is some quantum-mechanical operation, it must be linear (as we have shown that evolution in quantum mechanics is linear in general). Therefore:

$$\begin{aligned} C(|+\rangle \otimes |0\rangle) &= \frac{1}{\sqrt{2}} \left[C(|0\rangle \otimes |0\rangle) + C(|1\rangle \otimes |0\rangle) \right] \\ &= \frac{a}{\sqrt{2}} |0\rangle \otimes |0\rangle + \frac{b}{\sqrt{2}} |1\rangle \otimes |1\rangle \end{aligned} \tag{3.37}$$

where in the second equality we invoke Eq. (3.35), and a, b are the proportionality constants. However, the results in Eqs. , are not equal (or proportional); contradiction. Therefore C cannot exist. □

Note that the No-cloning theorem does not forbid cloning in a fixed basis - try for example constructing a quantum mechanical protocol that can clone $|\uparrow\rangle, |\downarrow\rangle$ states. What it does forbid is the cloning of arbitrary states, i.e. cloning in an arbitrary basis; this is where the linearity in the above proof kicks in to derive a contradiction. We leave the reader to ponder why the existence of classical copying machines is not in contradiction with the No-cloning theorem.

3.3 Superdense Coding and Quantum Teleportation

Now that we've looked at a non-application of entanglement, let's start to study some actual applications!

To set up our discussion; we introduce the notion of a quantum bit, or qubit, which generalizes the notion of a classical bit to a quantum setting. A bit is the fundamental unit of information classically, taking on one of two states, 0 or 1. A qubit as the quantum-mechanical fundamental unit of information can take on any complex superposition of the 0 and 1 states:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle. \tag{3.38}$$

$|0\rangle, |1\rangle$ are orthogonal basis states (like the classical bit) that span the Hilbert space $\mathcal{H} = \mathbb{C}^2$ that the qubit lives in. So really when we have been discussing spin-1/2 systems we have been talking about qubits all along! We have done nothing really more than re-label $|\uparrow\rangle$ with $|0\rangle$ and $|\downarrow\rangle$ with $|1\rangle$.

What is different compared to the classical setting is that qubits can be in superposition states, and they can be entangled with each other (if we have multiple of them). We now ask - what can we do with qubits? This question in generality is still a research field, with quantum algorithms being a exciting research area (some examples of which we will discuss later in this chapter). But to begin, maybe let's start with a slightly easier question; what is a qubit worth, relative to a classical bit?

As we have discussed before, to specify a qubit state, we have two complex coefficients α, β, but with the normalization constraint $\langle\psi|\psi\rangle = 1 \implies |\alpha|^2 + |\beta|^2 = 1$ and the irrelevance of global phase $|\psi\rangle \sim e^{i\varphi}|\psi\rangle$ a single qubit state is uniquely specified by two real numbers. To specify a real number, we require an infinite number of bits, so is the answer that a qubit is worth an infinite number of classical bits?

Well no, not quite; although a general qubit state is in fact specified by two real numbers, when we measure the qubit in some basis, we only ascertain one of two outcomes - in other words, we can only measure one bit worth of information from a qubit. So, is the answer that a qubit is worth exactly one classical bit?

The answer to this question is illuminated by the discussion of our first (real!) quantum protocol - known as superdense coding. We will see in this protocol that it is possible to encode *two* classical bits in a single qubit; provided we make use of entanglement.

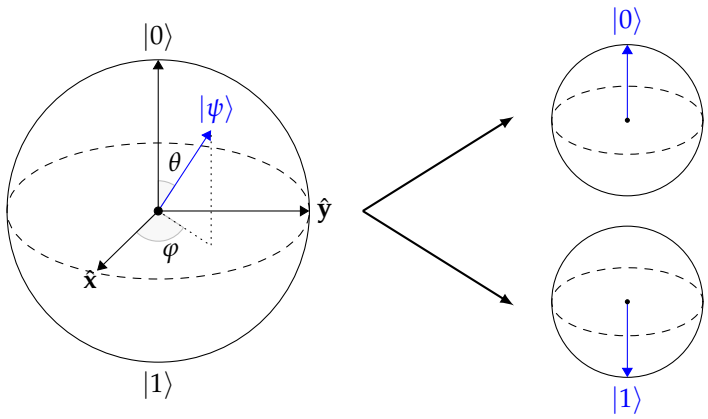

Figure 3.1: When we measure a qubit in the S_z eigenbasis (in quantum information lingo, called the computational basis) of $\{|0\rangle, |1\rangle\}$, we only find one of two outcomes, and the post measurement-state is one of $|0\rangle, |1\rangle$ - one of two states, just like the classical bit (this is true regardless of what single-qubit measurement basis we choose; the possible post-measurement states will be some two antipodal points on the Bloch sphere).

To discuss this protocol, we introduce the Bell basis - we have discussed the Bell state $|B_{11}\rangle$ already, but in fact there are four Bell states, which form a orthonormal basis (check!) for the Hilbert space of two qubits $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$:

$$\begin{aligned} |B_{00}\rangle &= \frac{|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle}{\sqrt{2}} \\ |B_{01}\rangle &= \frac{|0\rangle \otimes |0\rangle - |1\rangle \otimes |1\rangle}{\sqrt{2}} \\ |B_{10}\rangle &= \frac{|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle}{\sqrt{2}} \\ |B_{11}\rangle &= \frac{|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle}{\sqrt{2}} \end{aligned} \tag{3.39}$$

Recalling that $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ and $\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|$, we make the observation that the 01/10/11 Bell states are related to the 00 Bell state via the application of Paulis on the first qubit:

$$\begin{aligned} |B_{01}\rangle &= \sigma_{z,1}|B_{00}\rangle \\ |B_{10}\rangle &= \sigma_{x,1}|B_{00}\rangle \\ |B_{11}\rangle &= \sigma_{z,1}\sigma_{x,1}|B_{00}\rangle \end{aligned} \tag{3.40}$$

Which we may summarize with the relation:

$$|B_{ab}\rangle = (\sigma_{z,1})^b (\sigma_{x,1})^a |B_{00}\rangle. \tag{3.41}$$

Note that we could very well have applied the Paulis to the second qubit, i.e.:

$$|B_{ab}\rangle = (\sigma_{z,2})^b (\sigma_{x,2})^a |B_{00}\rangle. \tag{3.42}$$

Although we will not need it for the superdense coding protocol (we will need it for the following teleportation protocol), it will be useful to note one more property of $|B_{00}\rangle$, namely that it is the $+1$ eigenvalue

of $\sigma_z \otimes \sigma_z$ and $\sigma_x \otimes \sigma_x$ (check!):

$$\begin{aligned} \sigma_z \otimes \sigma_z |B_{00}\rangle &= |B_{00}\rangle \\ \sigma_x \otimes \sigma_x |B_{00}\rangle &= |B_{00}\rangle \end{aligned} \tag{3.43}$$

Physically, this means that S_z and S_x measurements on the two qubits of $|B_{00}\rangle$ are perfectly correlated. It is also worth noting that $|B_{00}\rangle$ is uniquely specified by the property that it is so-called *stabilized* by $\sigma_z \otimes \sigma_z$ and $\sigma_x \otimes \sigma_x$ - this is probably the first time you have seen states described in this manner, but if you go on to do more courses/research in quantum information theory (and in particular quantum error correction) you will see this method of specifying states (via the operators they are stabilized by) come up time and time again through the *stabilizer formalism*.

With this, we now have all the tools available to understand the superdense coding protocol, which we now lay out here.

Protocol: Superdense coding

Objective: Transmit two bits of information a, b from A to B.
Protocol:

1. In advance of the actual communication, the parties share a bell state $|B_{00}\rangle$ between themselves.
2. Sender A applies $(\sigma_z)^b(\sigma_x)^a$ to their qubit, encoding $(a, b) \in \mathbb{Z}_2 \times \mathbb{Z}_2$.
3. A sends their qubit to B.
4. B measures their qubit in the Bell basis[a]. Depending on the outcome, they recover a, b.

[a] Formally, they can measure some observable $O = \lambda_{00}|B_{00}\rangle\langle B_{00}| + \lambda_{01}|B_{01}\rangle\langle B_{01}| + \lambda_{10}|B_{10}\rangle\langle B_{10}| + \lambda_{11}|B_{11}\rangle\langle B_{11}|$, and since the state they have is one of the four Bell states, depending on which outcome λ_{ab} they measure they can recover the two bits a, b.

So given the above protocol, is the answer that one qubit is worth two classical bits? The answer is not so clear cut - we would not have been able to transmit two bits of information had we only sent over an unentangled qubit. Indeed, the entanglement here played a role, and in communicating the two bits of information, we have used up one bit of entanglement.

Let us also discuss the counterpart to the superdense coding protocol - namely, the quantum teleportation protocol. In superdense coding, we wanted to communicate two bits worth of information and so we physically sent a qubit; in the teleportion protocol things are reversed; we will communicate/teleport a qubit state, and in order to do so physically send two classical bits worth of information. The protocol is as follows.

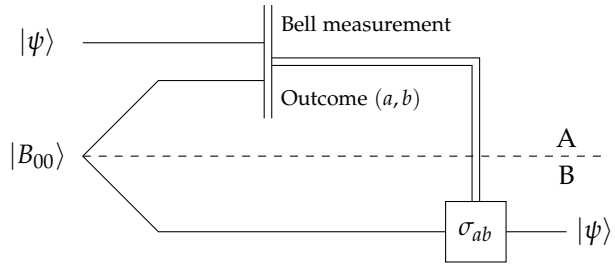

Figure 3.2: Graphical depiction of the quantum teleportation protocol.

Protocol: Quantum Teleportation

Objective: Transmit a qubit state $|\psi\rangle \in \mathbb{C}^2$ from A to B.
Protocol:

1. In advance of the actual communication, the parties share a bell state $|B_{00}\rangle$ between themselves.
2. A prepares the state $|\psi\rangle$ she wants to transmit (she now has two qubits; one from the shared Bell pair, and one for $|\psi\rangle$).
3. A performs a measurement in the Bell basis on their two qubits, and obtains the two-bit outcome (a, b).
4. A transmits the two-bit measurement outcome (a, b) to B.
5. B applies the correction operator $\sigma_{ab} = (\sigma_x)^a(\sigma_z)^b$ to their qubit. The resulting state of B's qubit is $|\psi\rangle$.

Proof of Correctness. Let $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ be the state that A wants to transmit to B. Let us label this qubit as qubit 1, A's half of the Bell pair as qubit 2, and B's half of the Bell pair as qubit 3. The initial state is then $|\psi\rangle_1 \otimes |B_{00}\rangle_{23}$. Now, we consider the action of the Bell measurement (with outcome (a, b)) on Alice's two qubits. Up to normalization, this has the effect of applying the projector:

$$\Pi_{ab,12} \otimes \mathbb{I}_3 = |B_{ab}\rangle\langle B_{ab}|_{12} \otimes \mathbb{I}_3 \tag{3.44}$$

onto the state. Using Eq. (3.40), we can write:

$$|B_{ab}\rangle_{ij} = (\sigma_{z,i})^b(\sigma_{x,i})^a|B_{00}\rangle_{ij} \tag{3.45}$$

and so we can rewrite the projector as:

$$\Pi_{ab,12} \otimes \mathbb{I}_3 = |B_{ab}\rangle\langle B_{00}|_{12}(\sigma_{x,2})^a(\sigma_{z,2})^b \otimes \mathbb{I}_3 \tag{3.46}$$

So applying this to the initial state, we find:

$$\begin{aligned}\Pi_{ab,12} \otimes I_3 \left(|\psi\rangle_1 \otimes |B_{00}\rangle_{23}\right) &= \left(|B_{ab}\rangle\langle B_{00}|_{12}(\sigma_{x,2})^a(\sigma_{z,2})^b \otimes \mathbb{I}_3\right) \left((\alpha|0\rangle_1 + \beta|1\rangle_1) \otimes |B_{00}\rangle_{23}\right) \\ &= \left(|B_{ab}\rangle\langle B_{00}|_{12} \otimes \mathbb{I}_3\right) \left((\alpha|0\rangle_1 + \beta|1\rangle_1) \otimes ((\sigma_{x,2})^a(\sigma_{z,2})^b \otimes \mathbb{I}_3)|B_{00}\rangle_{23}\right)\end{aligned} \tag{3.47}$$

Now using that $|B_{00}\rangle$ is stabilized by $\sigma_z \otimes \sigma_z$ and $\sigma_x \otimes \sigma_x$ (Equation (3.43)) we can write:

$$|B_{00}\rangle_{23} = (\sigma_{z,2} \otimes \sigma_{z,3})^b|B_{00}\rangle_{23} = (\sigma_{z,2} \otimes \sigma_{z,3})^b(\sigma_{x,2} \otimes \sigma_{x,3})^a|B_{00}\rangle_{23} \tag{3.48}$$

And now using that $\sigma_i^2 = \mathbb{I}$ for each of the Pauli matrices:

$$\begin{aligned}((\sigma_{x,2})^a(\sigma_{z,2})^b \otimes \mathbb{I}_3)|B_{00}\rangle_{23} &= ((\sigma_{x,2})^a(\sigma_{z,2})^b \otimes \mathbb{I}_3)(\sigma_{z,2} \otimes \sigma_{z,3})^b(\sigma_{x,2} \otimes \sigma_{x,3})^a|B_{00}\rangle_{23} \\ &= ((\sigma_{x,2})^a \otimes \mathbb{I}_3)((\sigma_{z,2})^2 \otimes \sigma_{z,3})^b(\sigma_{x,2} \otimes \sigma_{x,3})^a|B_{00}\rangle_{23} \\ &= ((\sigma_{x,2})^a \otimes \mathbb{I}_3)(\mathbb{I}_2 \otimes \sigma_{z,3})^b(\sigma_{x,2} \otimes \sigma_{x,3})^a|B_{00}\rangle_{23} \\ &= (\mathbb{I}_2 \otimes \sigma_{z,3})^b((\sigma_{x,2})^2 \otimes \sigma_{x,3})^a|B_{00}\rangle_{23} \\ &= (\mathbb{I}_2 \otimes \sigma_{z,3})^b(\mathbb{I}_2 \otimes \sigma_{x,3})^a|B_{00}\rangle_{23} \\ &= \mathbb{I}_2 \otimes (\sigma_{z,3})^b(\sigma_{x,3})^a|B_{00}\rangle_{23}\end{aligned} \tag{3.49}$$

So then Eq. (3.47) becomes:

$$\begin{aligned}\Pi_{ab,12}\otimes I_3\left(|\psi\rangle_1\otimes|B_{00}\rangle_{23}\right) &= (\sigma_{z,3})^b(\sigma_{x,3})^a\left(|B_{ab}\rangle\langle B_{00}|_{12}\otimes\mathbb{I}_3\right)\left(|\psi\rangle_1\otimes|B_{00}\rangle_{23}\right)\\ &= (\sigma_{z,3})^b(\sigma_{x,3})^a\left(\frac{|B_{ab}\rangle\langle 00|_{12}+|B_{ab}\rangle\langle 11|_{12}}{\sqrt{2}}\otimes\mathbb{I}_3\right)\left(\frac{\alpha|000\rangle_{123}+\beta|100\rangle_{123}+\alpha|011\rangle_{123}+\beta|111\rangle_{123}}{\sqrt{2}}\right.\\ &= (\sigma_{z,3})^b(\sigma_{x,3})^a\frac{1}{2}\left(|B_{ab}\rangle_{12}\right)\otimes(\alpha|0\rangle_3+\beta|1\rangle_3)\\ &= \frac{1}{2}|B_{ab}\rangle_{12}\otimes\left((\sigma_{z,3})^b(\sigma_{x,3})^a|\psi\rangle_3\right)\end{aligned} \tag{3.50}$$

So indeed we have teleported $|\psi\rangle$ to the third (B's) qubit, up to applying a correction operator of $\sigma_{ab} = (\sigma_{x,3})^a(\sigma_{z,3})^b$.

□

Note that although the name might suggest some kind of superluminal communication, the teleportation protocol is fully consistent with special relativity - in order to recover the correct state $|\psi\rangle$ at the end, A must transmit to B the two-bit measurement outcome of the Bell basis measurement; this communication cannot exceed light speed.

A tangent to conclude this section that is certainly beyond the scope of this course but nevertheless interesting; a variant of the teleportation protocol (half-teleportation) forms the backbone of measurement-based quantum computation, where a computation is carried out solely via a sequence of local (adaptive) measurements on an initial resource state. You can read more about it here, among other places.

3.4 Density Operators

At the beginning of the course, we introduced the state ket in a Hilbert space $|\psi\rangle \in \mathcal{H}$ which represents a quantum state, and how it was a more abstract/general object than the wavefunction (which is nothing but the coefficients of the state ket expanded in a particular basis) $\psi(x)$ which was the primary objects of study in your first quantum mechanics course. However, it turns out that state kets are not the most general way of representing quantum states - the most general representation turns out to be density operators, and we give two examples (based on material already discussed in the course) as motivation for them.

The first example comes from the Stern-Gerlach experiment. Consider a stream of spin-1/2 particles where each spin is in the spin-up state with probability 1/2 and spin-down state with probability 1/2.How would we represent this physical situation?

Because it has a 50/50 probability of being measured to be up or down, we might consider the equal weight superposition of the $|\uparrow\rangle, |\downarrow\rangle$ states:

$$|\psi\rangle \stackrel{?}{=} \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}} \tag{3.51}$$

We note that this is of course just the S_x eigenstate of $|\rightarrow\rangle$. However, although this state indeed reproduces the correct measurement statistics for an S_z measurement (where we get 50/50 outcomes), it must give the correct probabilities for any possible observable on the given physical system. To this end, we consider a measurement of S_x. By the laws of conditional probability, the probability of measuring spin right for a beam which has 50% probability of being spin up or down should be:

$$p(\rightarrow) = p(\uparrow)p(\rightarrow \mid \uparrow) + p(\downarrow)p(\rightarrow \mid \downarrow) = \frac{1}{2}|\langle\rightarrow|\uparrow\rangle|^2 + \frac{1}{2}|\langle\leftarrow|\downarrow\rangle|^2 = \frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2} \tag{3.52}$$

In the above, $p(\uparrow\ /\ \downarrow)$ denotes the probability that the spin we draw from the beam is spin up or spin down, and $p(\rightarrow \mid \uparrow\ /\ \downarrow)$ denotes the probability that we measure spin-right given an input state of spin-up/spin-down. So, for our physical scenario, we have a probability of measuring spin-right of $\frac{1}{2}$. However

for our guess of $|\psi\rangle = |\rightarrow\rangle$, of course $p(\rightarrow) = 1$ and so our guess is incorrect. We might try to refine our guess, but any such attempt will be futile. For the provided physical situation of the 50/50 up-down beam, we will have a 50/50 distribution of outcome no matter how we orient the Stern-Gerlach measurement appratus (check!) - however any state ket $|\psi\rangle$ will have to be an eigenstate of spin in some direction, and therefore will not be able to reproduce the measurement statistics for all possible observables. So, there is something that the state ket $|\psi\rangle$ cannot quite capture about this scenario.

The second example comes from our thought experiment of the entangled Bell pair, where one half of the Bell pair is taken to the moon while the other remains on Earth. Suppose I was the experimenter on Earth, and I want some description of the qubit on Earth only, from which I can derive all statistics of the measurements that I can do on it. In the state ket formalism, the fact that the state is entangled means that (by definition) I cannot associate any state ket to the Earth qubit alone. So we require a more general formalism to handle this scenario, as well[19].

Before jumping to what this generalization is, we first require one piece of mathematical machinery - namely, the *trace* operation.

Definition: Trace

Let A be any operator acting on a Hilbert space $\mathcal{H}$. The trace of A is then defined as:

$$\mathrm{Tr}(A) := \sum_i \langle i|A|i\rangle \tag{3.53}$$

where the sum is taken over any ONB $\mathcal{B} = \left\{|i\rangle\right\}_{i=1}^{\dim(\mathcal{H})}$.

Proposition: Properties of the trace

1. The trace of A is independent of the ONB chosen to evaluate it.
2. The trace of A is equal to the sum of eigenvalues of A.
3. The trace is cyclic, i.e. for operators A, B, C:

$$\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB). \tag{3.54}$$

Proof. 1. Let $\left\{|i\rangle\right\}_i, \left\{|j\rangle\right\}_j$ be two ONBs. We can express the $|j\rangle$s in terms of the $|i\rangle$s by:

$$|j\rangle = \sum_i \langle i|j\rangle |i\rangle \tag{3.55}$$

[19] Although taking a qubit to the moon probably isn't of interest to most experimentalists, a more experimentally relevant and related scenario is when considering some system of interest and a larger, uncontrollable environment it is entangled with. In this case, we would want some way of just describing the system and deriving all measurement statistics of the system, somehow ignoring the larger environment it is coupled to.

Taking the trace using the $\{|j\rangle\}_j$ ONB, we have:

$$\begin{aligned}
\sum_j \langle j|A|j\rangle &= \sum_j \left(\sum_i \langle j|i\rangle\langle i|\right) A \left(\sum_{i'} \langle i'|j\rangle |i'\rangle\right) \\
&= \sum_{i,i'} \left(\langle i'| \left(\sum_j |j\rangle\langle j|\right) |i\rangle\right) \langle i|A|i'\rangle \\
&= \sum_{i,i'} \left(\langle i'|\mathbb{I}|i\rangle\right) \langle i|A|i'\rangle \\
&= \sum_{i,i'} \langle i'|i\rangle\langle i|A|i'\rangle \\
&= \sum_{i,i'} \delta_{ii'}\langle i|A|i'\rangle \\
&= \sum_i \langle i|A|i\rangle
\end{aligned} \tag{3.56}$$

where in the third equality we use the resolution of the identity. Hence the trace evaluated using the two ONBs are equivalent, proving the claim.

2. Take $\mathcal{B}$ to be the eigenbasis of A; the claim immediately follows.

3. First, we show that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ for two operators A, B. Let $C = AB$ and $D = BA$. Consider some ONB $\{|i\rangle\}_i$. Inserting the resolution of the identity, we have:

$$\begin{aligned}
\langle i|C|j\rangle &= \sum_m \langle i|A|m\rangle\langle m|B|j\rangle \\
\langle i|D|j\rangle &= \sum_m \langle i|B|m\rangle\langle m|A|j\rangle
\end{aligned} \tag{3.57}$$

And so taking the trace of C:

$$\begin{aligned}
\mathrm{Tr}(C) &= \sum_i \langle i|C|i\rangle \\
&= \sum_i \sum_m \langle i|A|m\rangle\langle m|B|i\rangle \\
&= \sum_i \sum_m \langle m|B|i\rangle\langle i|A|m\rangle \\
&= \sum_m \sum_i \langle m|B|i\rangle\langle i|A|m\rangle \\
&= \sum_m \langle m|D|m\rangle \\
&= \mathrm{Tr}(D)
\end{aligned} \tag{3.58}$$

So $\mathrm{Tr}(C) = \mathrm{Tr}(D)$. To prove $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA)$, take $A \to A, B \to BC$ in the above, and to prove $\mathrm{Tr}(BCA) = \mathrm{Tr}(CAB)$, take $A \to B, B \to CA$ in the above.

□

Now that we've discussed the trace, we can move on to constructing and defining the density operator. Consider an ensemble $\mathcal{E}$ of states $|\varphi_i\rangle$ drawn rnadomly, with probabilities p_i (e.g. in the Stern-Gerlach motivating example, drawing $|\uparrow\rangle$ with $p_\uparrow = \frac{1}{2}$ and $|\downarrow\rangle$ with $p_\downarrow = \frac{1}{2}$)"

$$\mathcal{E} = \{(p_i, |\varphi_i\rangle)\}_i \tag{3.59}$$

Given $\mathcal{E}$, for any operator A, the expectation $\langle A \rangle_{\mathcal{E}}$ is:

$$\langle A \rangle_{\mathcal{E}} = \sum_i p_i \langle \varphi_i | A | \varphi_i \rangle \tag{3.60}$$

i.e. the sum of expectation values for the states in the ensemble, weighted by the probability of drawing them. Let us manipulate this expression by inserting the resolution of the identity:

$$\begin{aligned}
\langle A \rangle_{\mathcal{E}} &= \sum_i p_i \langle \varphi_i | A | \varphi_i \rangle \\
&= \sum_i p_i \sum_j \langle \varphi_i | j \rangle \langle j | A | \varphi_i \rangle \\
&= \sum_i p_i \sum_j \langle j | A | \varphi_i \rangle \langle \varphi_i | j \rangle \\
&= \sum_i p_i \mathrm{Tr}(A | \varphi_i \rangle \langle \varphi_i |) \\
&= \mathrm{Tr}(A \left[\sum_i p_i | \varphi_i \rangle \langle \varphi_i | \right]) \\
&= \mathrm{Tr}(A\rho)
\end{aligned} \tag{3.61}$$

where in the last equality we define the density operator $\rho = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|$. The density operator corresponding to an ensemble contains all the physically relevant information that the ensemble contains (e.g. all relevant statistics of measurement and evolution can be done with the density operator). It is worth noting however that a given density operator does not correspond to a unique ensemble (can you show this by example?)

Definition: Density Operators

For any ensemble $\mathcal{E} = \{(p_i, |\varphi_i\rangle)\}_i$,

$$\rho \coloneqq \sum_i p_i |\varphi_i\rangle\langle\varphi_i| \tag{3.62}$$

is the corresponding density operator.

Proposition: Density Operator Properties

For all density operators ρ it holds that:

1. $\mathrm{Tr}(\rho) = 1$ (Normalization)
2. $\rho^\dagger = \rho$ (Hermicity)
3. $\langle \psi | \rho | \psi \rangle \geq 0, \quad \forall |\psi\rangle \in \mathcal{H}$ (Positivity)

Proof. Left as a homework exercise. □

Definition: Pure and Mixed States

If a density operator ρ has a single eigenvalue of 1 and all other eigenvalues 0, it is called pure, and $\rho = |\varphi\rangle\langle\varphi|$. Otherwise, it is called mixed.

Note that pure operators $\rho = |\varphi\rangle\langle\varphi|$ correspond to state kets $|\varphi\rangle$. However, the density operator formalism is able to account for probabilistic mixtures (mixed states) which was not something that state kets were able to represent.

Proposition: Criterion for purity

All density operators ρ satisfy $\mathrm{Tr}(\rho^2) \leq 1$, with equality if and only if ρ is pure.

Proof. Left as homework exercise. □

Axiom: Quantum measurement (with density operators)

Consider a measurement with corresponding projectors $\mathcal{M} = \{\Pi_i, i = 1, \ldots k\}$ on a density operator ρ.
Dirac postulate: When we measure outcome i, the density operator evolves as:

$$\rho \mapsto \frac{\Pi_i \rho \Pi_i}{\mathrm{Tr}(\Pi_i \rho)} \tag{3.63}$$

Born rule: The probability of measuring outcome i is given by:

$$p(i) = \mathrm{Tr}(\Pi_i \rho). \tag{3.64}$$

Axiom: Unitary evolution (with density operators)

Considering the time-dependent Schrödinger equation as a PDE for time-evolution operators:

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = H U(t, t_0) \tag{3.65}$$

with $U(t_0, t_0) = \mathbb{I}$. Density operators evolve as:

$$\rho(t) = U(t, t_0) \rho(t_0) U^\dagger(t, t_0) \tag{3.66}$$

or expressed in terms of a PDE:

$$i\hbar \frac{\partial}{\partial t} \rho(t) = [H, \rho(t)]. \tag{3.67}$$

You can verify that when ρ is a pure state, the Dirac postulate/Born rule/Unitary evolution stated in the density operator formalism reduces to the familiar axioms for measurement/unitary evolution of state kets.

To close out this section, we return to the motivating example; how do we represent the state corresponding to a beam of spin-1/2 particles where each spin in the beam is spin-up with probability 1/2 and spin-down with probability 1/2? Note that this scenario is precisely what you encountered in the first homework assignment; there, we fed a beam of spin-right particles into a Stern-Gerlach apparatus that measured the z-component of spin (so by the Born rule there is a 50/50 distribution of outcome for measuring spin-up/spin-down) and then we recombined the two up/down beams such that we "forgot" the measurement outcome for a given individual spin in the final combined beam.

The answer is that we have a mixed state; the ensemble corresponding to the beam in this scenario is

$\left\{(\frac{1}{2}, |\uparrow\rangle), (\frac{1}{2}, |\downarrow\rangle)\right\}$. If we then calculate the corresponding density operator, we find:

$$\rho_{50/50} = \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\downarrow\rangle\langle\downarrow| \cong \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.68}$$

Compare this to the density operator corresponding to the incoming beam, which is a pure state:

$$\rho_{\rightarrow} = |\rightarrow\rangle\langle\rightarrow| \cong \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{3.69}$$

Note that in both cases we have represented the density operator in the S_z eigenbasis, under the identification $|\uparrow\rangle \cong \binom{1}{0}$ and $|\downarrow\rangle \cong \binom{0}{1}$. We make a key observation - in the case of $\rho_{\rightarrow}$, we have non-vanishing off-diagonal entries. The meaning of these off-diagonal elements is the coherences in between the two basis states we are expanding in. In the case of $\rho_{\rightarrow}$, we have a coherent superposition between $|\uparrow\rangle, |\downarrow\rangle$ (namely - $|+\rangle = \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}$) and so the off-diagonals have maximal magnitude. In the case of $\rho_{50/50}$, the off-diagonals are zero[20], indicating that we have a decoherent mixture of the two basis states $|\uparrow\rangle, |\downarrow\rangle$.

In conclusion, let us quickly summarize what the meaning of the density matrix elements are (when represented in some ONB $\mathcal{B} = \{|i\rangle\}_i$). The diagonal entries $[\rho]_{ii} = \langle i|\rho|i\rangle$ tells us about the probability of obtaining the basis state $|i\rangle$ if we conduct a measurement of ρ in the basis $\mathcal{B}$. The diagonal entries $[\rho]_{ij} = \langle i|\rho|j\rangle$ give us information about the coherences between the states $|i\rangle, |j\rangle$ in the basis.

3.5 Reduced Density Operators and Impossibility of Superluminal Communication

We've seen how density operators can handle probabilistic (decoherent) mixtures; in this section we will see how they are able to handle describing some portion of a larger (potentially entangled) state, as in the Bell state between Earth and Moon scenario.

We will require one more piece of mathematical machinery before jumping into it though; namely, the *partial trace*.

Definition: Partial trace

Consider an operator X_{AB} acting on a composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Let $\mathcal{B}_A = \{|i\rangle, i = 1, \ldots, d_A\}, \mathcal{B}_B = \{|j\rangle, j = 1, \ldots, d_B\}$ be ONBs on A/B. Then, the partial trace $\mathrm{Tr}_A(X_{AB})$ is defined as:

$$\mathrm{Tr}_A(X_{AB}) := \sum_i^{d_A} {}_A\langle i|X_{AB}|i\rangle_A \tag{3.70}$$

where the sum is taken over $\mathcal{B}_A$. $\mathrm{Tr}_B(X_{AB})$ is defined analogously.

This definition requires some parsing - it is not immediately clear how an operator X_{AB} which acts on the composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ acts on a basis ket $|i\rangle$ of $\mathcal{H}_A$. To this end, we note that the pairwise tensor products of states from $\mathcal{B}_A$ and $\mathcal{B}_B$ form a basis for $\mathcal{H}_A \otimes \mathcal{H}_B$. So, we can insert two resolutions of the identity:

$$X_{AB} = \sum_{k,r}^{d_A}\sum_{l,s}^{d_B} |k\rangle_A\langle k| \otimes |l\rangle_B\langle l|X_{AB}|r\rangle_A\langle r| \otimes |s\rangle_B\langle s| \tag{3.71}$$

We also observe that:

$${}_A\langle i|\left(|k\rangle_A \otimes |l\rangle_B\right) = \langle i|k\rangle_A|l\rangle_B = \delta_{ik}|l\rangle_B \tag{3.72}$$

[20] Some terminology - in the cases where $\rho \propto \mathbb{I}$, we say that we have a maximally mixed state. In this case there is no coherence between the basis states whatsoever.

Therefore, with respect to these bases:

$$\begin{aligned}\mathrm{Tr}_A(X_{AB}) &= \sum_i^{d_A} {}_A\langle i| \left(\sum_{k,r}^{d_A} \sum_{l,s}^{d_B} |k\rangle_A\langle k| \otimes |l\rangle_B\langle l|X_{AB}|r\rangle_A\langle r| \otimes |s\rangle_B\langle s| \right) |i\rangle_A \\ &= \sum_{i,k,r}^{d_A} \sum_{l,s}^{d_B} \delta_{ik}|l\rangle_B \left({}_A\langle k| \otimes {}_B\langle l|X_{AB}|r\rangle_A \otimes |s\rangle_B \right) {}_B\langle s|\delta_{ir} \\ &= \sum_i^{d_A} \sum_{l,s}^{d_B} |l\rangle_B \left({}_A\langle i| \otimes {}_B\langle l|X_{AB}|i\rangle_A \otimes |s\rangle_B \right) {}_B\langle s|\end{aligned} \tag{3.73}$$

Note that partial traces have the following important property (which we state without proof, but if you are able to parse the definition of the partial trace, it should follow from the definition without too much trouble):

Proposition: Trace and partial trace

Let X_{AB} be an operator on the composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Then:

$$\mathrm{Tr}(X) = \mathrm{Tr}_B(\mathrm{Tr}_A(X_{AB})) = \mathrm{Tr}_A(\mathrm{Tr}_B(X_{AB})). \tag{3.74}$$

We also note that the partial trace, like the trace has a cyclicity property (although weaker):

Proposition: Cyclicity of partial trace

Let X_{AB} be an operator acting on a composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ and let T_A, S_A be operators acting on $\mathcal{H}_A$. Then:

$$\mathrm{Tr}_A([T_A \otimes I_B]X_{AB}[S_A \otimes I_B]) = \mathrm{Tr}_A([S_A T_A \otimes I_B]X_{AB}). \tag{3.75}$$

We now construct the *reduced density operator*. Consider any observable $O_A \otimes I_B$ (i.e. where we have some observable of interest on A only - you can think of A as the Earth and B as the moon in the Earth-Moon Bell state scenario). Consider an arbitrary bipartite state $|\psi\rangle_{AB}$. Calculating the expectation value of $O_A \otimes I_B$ with respect to this state, we have:

$$\begin{aligned}\langle O_A \otimes I_B \rangle_\psi &= \mathrm{Tr}(O_A \otimes I_B |\psi\rangle_{AB}\langle\psi|) \\ &= \mathrm{Tr}_A(\mathrm{Tr}_B(O_A \otimes I_B |\psi\rangle_{AB}\langle\psi|)) \\ &= \mathrm{Tr}_A(O_A \mathrm{Tr}_B(|\psi\rangle_{AB}\langle\psi|)) \\ &= \mathrm{Tr}_A(O_A \rho_A)\end{aligned} \tag{3.76}$$

where in the first equality we take the expectation value in the density operator formalism (with $\rho_{AB} = |\psi\rangle_{AB}\langle\psi|$), in the second equality we use that $\mathrm{Tr}(\cdot) = \mathrm{Tr}_A(\mathrm{Tr}_B(\cdot))$ as in the above proposition, and in the final equality we define $\rho_A \coloneqq \mathrm{Tr}_B(|\psi\rangle_{AB}\langle\psi|)$ as the reduced density operator on subsystem A.

Definition: Reduced density operator

Given a quantum state (density operator) ρ_{AB} acting on a composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, the reduced density operator on subsystem A is given by:

$$\rho_A \coloneqq \mathrm{Tr}_B(\rho_{AB}) \tag{3.77}$$

By tracing out the B subsystem, we are able to obtain an object (the reduced density operator - which itself is a density operator) that contains all information about the statistics we would obtain from operations we perform on subsystem A only. Note that (crucially) this formalism still applies even if the joint quantum state ρ_{AB} is entangled (that is, $\rho_{AB} \neq \rho_1 \otimes \rho_2$ for any ρ_1, ρ_2).

Using the formalism of reduced density operators, we are able to prove the following:

Theorem: No superluminal communication

Quantum mechanics does not provide for superluminal communication.

We previously ruled out one proposed protocol on the basis of the no-cloning theorem; here we prove that any such proposed protocol is destined to fail.

Proof. Consider any state ρ_{AB} shared between parties A and B. We prove that superluminal communication is impossible by showing that it is impossible for a local operation at A to change the reduced density operator on subsystem B:

$$\rho_B := \mathrm{Tr}_A(\rho_{AB}). \tag{3.78}$$

In order to do this, we show that no form of quantum mechanical evolution on subsystem A can change ρ_B; in other words, we show that it is (i) impossible for local unitary evolution on subsystem A to change ρ_B and (ii) local measurement on subsystem A to change ρ_B.

(i) (Unitary Evolution). Unitary evolution on A evolves ρ_{AB} as:

$$\rho_{AB} \mapsto (U_A \otimes I_B)\rho_{AB}(U_A^\dagger \otimes I_B) \tag{3.79}$$

But we observe:

$$\mathrm{Tr}_A((U_A \otimes I_B)\rho_{AB}(U_A^\dagger \otimes I_B)) = \mathrm{Tr}_A((U_A^\dagger \otimes I_B)(U_A \otimes I_B)\rho_{AB}) = \mathrm{Tr}_A(\rho_{AB}) = \rho_B \tag{3.80}$$

where we use the cyclicity of the partial trace. We conclude that ρ_B is unchanged by local unitary evolution on subsystem A.

(ii) (Measurement). We consider a measurement on A from the perspective of B. Since B doesn't know the outcome of the measurement, ρ_{AB} is replaced with the average over all possible measurement outcomes:

$$\rho_{AB} \mapsto \sum_i \frac{(\Pi_{i,A} \otimes I_B)\rho_{AB}(\Pi_{i,A} \otimes I_B)}{\mathrm{Tr}(\Pi_{i,A}\rho_{AB})}\mathrm{Tr}(\Pi_{i,A}\rho_{AB}) = \sum_i (\Pi_{i,A} \otimes I_B)\rho_{AB}(\Pi_{i,A} \otimes I_B) \tag{3.81}$$

where the $\frac{(\Pi_{i,A} \otimes I_B)\rho_{AB}(\Pi_{i,A} \otimes I_B)}{\mathrm{Tr}(\Pi_{i,A}\rho_{AB})}$ comes from the Dirac postulate in the density operator formalism, and the $\mathrm{Tr}(\Pi_{i,A}\rho_{AB})$ is the weight in the average. Now tracing out subsystem A, we have:

$$\mathrm{Tr}_A(\sum_i (\Pi_{i,A} \otimes I_B)\rho_{AB}(\Pi_{i,A} \otimes I_B)) = \mathrm{Tr}_A(\sum_i \Pi_{i,A}^2 \otimes I_B\rho_{AB}) = \mathrm{Tr}_A(\rho_{AB}) = \rho_B \tag{3.82}$$

where we have again used the cyclicity of the partial trace, as well as the completeness relation for projectors. We conclude that local measurement on subsystem A does not change ρ_B.

□

3.6 Quantum Cryptography

There are many instances (e.g. protecting your credit card information for an online transaction) in which we require secure communication - where the information is only accessible to the sender and recipient, and inaccessible to potential eavesdroppers. On the end of cryptographic protocols, we have two desiredata:

(a) The protocols are *unconditionally secure* - that is, the security of the protocol is based on the laws of physics/information theory.

(b) The protocols are *practical* - that is, they can be carried out via local operations and (potentially quantum) communication.

	Impractical	Practical
Secure	Classical	Quantum
Insecure	(Undesirable)	Classical

Table 3: Feasibility of cryptography using classical/quantum resources. With classical cryptography, we have secure but impractical protocols, and practical but insecure protocols. Quantum resources are able to provide protocols which are both secure and practical.

Before moving into quantum protocols, we look at some classical counterparts.

Protocol: One-time pad

Objective: To (securely) send a bitstring $\mathbf{x}$ from A to B.
Protocol:

1. The communicating parties A and B share a random bit string $\mathbf{r}$ (for simplicitly - the message $\mathbf{x}$ and the bitstring $\mathbf{r}$ contain the same number of bits), e.g.:

$$\mathbf{r} = (0001010100101001010001110 0 \ldots) \tag{3.83}$$

That is, A holds a copy of $\mathbf{r}$ and B does too.

2. A computes:

$$\mathbf{c} := \mathbf{x} \oplus \mathbf{r} \tag{3.84}$$

where $\oplus$ denotes component-wise addition modulo 2 (i.e. $0+0=1+1=0, 0+1=1+0=1$ for each component of the bitstring), and broadcasts $\mathbf{c}$ to the world.

3. B receives the encoded message $\mathbf{c}$, and computes:

$$\mathbf{c} \oplus \mathbf{r} = \mathbf{x} \oplus \mathbf{r} \oplus \mathbf{r} = \mathbf{x}. \tag{3.85}$$

The message $\mathbf{x}$ is thus successfully transmitted.

The one-time pad is secure, but impractical - the impracticality arises from the necessity of having a shared random bitstring $\mathbf{r}$. Of course one cannot repeatedly use the same bitstring to encode information, as then the code could be broken (hence - the "one-time" pad). Therefore, at some point one runs of randomly shared bitstring, and information cannot be communicated securely any longer.

We now consider a classical protocol in the other quadrant of Table 3.

Protocol: RSA (Sketch)

Objective: To (securely) send a bitstring $\mathbf{x}$ from A to B.
Protocol:

1. B chooses two prime numbers m, n and broadcasts $m \cdot n$.
2. A uses $m \cdot n$ as a key to encode the message $\mathbf{x}$, and broadcasts $m \cdot n(\mathbf{x})$.
3. B receives and decodes the encoded message $m \cdot n(\mathbf{x})$ using m, n.

The full protocol in detail can be found here.

Why is this protocol not unconditionally secure? It is because it rests on the computational assumption that factoring $m \cdot n$ into the prime factors m, n is computationally hard - it is widely believed that there exists no polynomial time algorithm for prime factorization[21]; however this has not been proven. There is no physical or information theoretic law guaranteeing its security.

We now discuss a quantum cryptographic protocol which is both unconditionally secure and practical.

Protocol: BB84

Objective: To (securely) send a bitstring $\mathbf{x}$ from A to B.
Protocol:

1. B the communicating parties A, B share a large number of Bell states $|B_{00}\rangle^{\otimes n}$.
2. A and B randomly measure their halves of Bell pairs either in the Z or X basis.
3. A and B make public their choices of measurement bases (not outcomes).
4. The measurement outcomes on the qubit pairs where the measurement bases agreed form the key. Of this key, B publishes a small subset of bits.
5. A compares the measurement outcomes published by B to her own. If there is perfect agreement, the remainder of the key is kept. Otherwise, the protocol is aborted (Note: not establishing communication is not a failure, only being eavesdropped is).
6. The key obtained by the above can now be used to send $\mathbf{x}$ from A to B securely, using the one-time pad protocol.

Let's study each step of this protocol. In the first step, we have the sharing of a large number of Bell states $|B_{00}\rangle_{AB} = \frac{|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B}{\sqrt{2}}$. Why these states in particular? We recall the property (Eq. (3.43)) that Z and X measurements are perfectly correlated[22] for $|B_{00}\rangle$. So, if we do an X measurement, we have the post measurement states:

$$\begin{aligned} \text{outcome "+": } |B_{00}\rangle_{AB} &\to |+\rangle_A \otimes |+\rangle_B \\ \text{outcome "-": } |B_{00}\rangle_{AB} &\to |-\rangle_A \otimes |-\rangle_B \end{aligned} \tag{3.86}$$

where $|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$ are the usual S_x eigenstates (with some new notation, commonly used in the quantum

[21] at least classically - Shor's celebrated quantum algorithm is a polynomial time algorithm for prime factorization, but it requires a quantum computer. So, quantum computers are able to break RSA. Stay tuned for the end of the chapter for some discussion.

[22] Since $|B_{11}\rangle$ displays strict anti-correlation of measurement outcomes in any local measurement basis, one might ask then whether $|B_{00}\rangle$ displays strict correlation of measurement outcomes in every local measurement basis. The answer turns out to be no - moreover, no bipartite state can display perfect correlation in every local measurement basis. Can you show it?

information literature). If we do an Z measurement, we instead get the outcomes:

$$\begin{aligned} \text{outcome "+": } |B_{00}\rangle_{AB} &\to |0\rangle_A \otimes |0\rangle_B \\ \text{outcome "-": } |B_{00}\rangle_{AB} &\to |1\rangle_A \otimes |1\rangle_B \end{aligned} \tag{3.87}$$

So, in the second step where Alice/Bob measure their halves of Bell pairs in the Z or X basis, if the measurement bases agree then they will see strict correlation of measurement outcomes, and if the measurement bases disagree then they will see no correlation.

We give an example of steps 3-5 in Table 4 below.

Bases	1	2	3	4	5	6	7	8	9
A	X	Z	X	Z	Z	Z	X	Z	X
B	Z	Z	X	Z	Z	X	X	X	Z
Outcomes	1	2	3	4	5	6	7	8	9
A	+	-	+	+	-	-	+	-	+
B	+	-	+	+	-	+	+	-	-

Table 4: A and B publicly publish their measurement bases (step 3), and see which ones agree (green) - the bell pairs for which the measurement bases disagree (red) are discarded. B then publishes a small subset of his measurement results (here, results -+ from 2/7, blue) for which the bases agreed (step 4). If A and B agree on the measurement outcomes for those, then the key consisting of the remaining measurement outcomes for which the bases agree (here ++- (001) from 3/4/5, pink) is kept and can be used for a one-time pad protocol.

The protocol works as if is no eavesdropping, the measurement outcomes for which the measurement bases are the same should agree (and this is checked via B publishing a subset of his measurement outcomes for which the bases are the same), and then the remainder of the outcomes is unknown to the world (only the bases are known) while A/B share it in confidence.

Let us explore one strategy that an eavesdropper E might take to eavesdrop on the key, and show why the strategy fails. In step 1 of the protocol, suppose E intercepts the Bell states shared between A and B and entangles herself with them:

$$|B_{00}\rangle_{AB} = \frac{|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B}{\sqrt{2}} \mapsto \frac{|0\rangle_A \otimes |0\rangle_B \otimes |0\rangle_E + |1\rangle_A \otimes |1\rangle_B \otimes |1\rangle_E}{\sqrt{2}} =: |GHZ\rangle_{ABE} \tag{3.88}$$

where the rightmost state is what is known as the Greenberger-Horne-Zeilinger state. Now, if A and B measure in the Z basis ($\{|0\rangle, |1\rangle\}$), we have:

$$\begin{aligned} \text{outcome "+": } |GHZ\rangle_{ABE} &\to |0\rangle_A \otimes |0\rangle_B \otimes |0\rangle_E \\ \text{outcome "-": } |GHZ\rangle_{ABE} &\to |1\rangle_A \otimes |1\rangle_B \otimes |1\rangle_E \end{aligned} \tag{3.89}$$

So E learns of the measurement outcome, and therefore the key! However, if A and B measure in the X basis ($\{|+\rangle, |-\rangle\}$), what happens? Let's work it out. First, the full density operator corresponding to the GHZ state is:

$$\rho_{ABE} = |GHZ\rangle_{ABE}\langle GHZ| = \frac{|000\rangle_{ABE}\langle 000| + |000\rangle_{ABE}\langle 111| + |111\rangle_{ABE}\langle 000| + |111\rangle_{ABE}\langle 111|}{2} \tag{3.90}$$

so tracing over Eve's qubit:

$$\rho_{AB} = \mathrm{Tr}_E(\rho_{ABC}) = {}_E\langle 0|\rho_{ABE}|0\rangle_E + {}_E\langle 1|\rho_{ABE}|1\rangle_E = \frac{|00\rangle_{AB}\langle 00| + |11\rangle_{AB}\langle 11|}{2} \tag{3.91}$$

And so if we calculate $\langle \sigma_{x,A} \otimes \sigma_{x,B} \rangle$, we find:

$$\langle \sigma_{x,A} \otimes \sigma_{x,B} \rangle = \mathrm{Tr}((\sigma_{x,A} \otimes \sigma_{x,B})\rho_{AB}) = \mathrm{Tr}(\frac{|00\rangle_{AB}\langle 11| + |11\rangle_{AB}\langle 00|}{2}) = 0. \tag{3.92}$$

So, we find that X-measurements of A and B are uncorrelated! However, A and B know that if they both measure the bell state $|B_{00}\rangle$ in the X-basis that the measurement outcomes should be perfectly correlated. Therefore, when they find that the measurements are uncorrelated (in step 5 of the protocol) Eve would be detected, and the protocol would be aborted.

Specifically: for each Bell state $|B_{00}\rangle$ that E entangles herself with (that gets used in step 4; i.e. for each Bell state that A/B measure in the same basis for, and for which B publishes the measurement outcomes publicly), she learns a bit of information with probability $\frac{1}{2}$ (she learns 1 bit if A/B measure in Z, and nothing if A/B measure in X). So, for each Bell pair, E can be said to learn half a bit of information. What is the probability of E being detected? Well, if A/B measure in the Z basis (which occurs with probability $\frac{1}{2}$ since they pick randomly from Z/X), then E is undetected. If A/B measure in the X basis (probability $1/2$) then the measurement outcomes are uncorrelated; so 50% of the time they will measure the same outcome and not detect E eavesdropping, and 50% of the time they measure different outcomes and detect E eavesdropping. So, for each Bell pair, the probability of E being detected is $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

In other words, for each half bit of information E, there is probability $\frac{1}{4}$ for E to be detected. So, if E learns n bit of information, the probability they are not detected is $\left(\frac{3}{4}\right)^{2n} = \left(\frac{9}{16}\right)^{n}$ so the probability of E not being detected decays exponentially in the number of bits that they learn. So, the protocol is secure against this kind of eavesdropping. This concludes our discussion of perhaps the most accessible near-term application of quantum entanglement. Time permitting, we will discuss another application - quantum computation - at the end of the course.