

PHYS 402 (Applications of Quantum Mechanics) Notes

Rio Weil

This document was typeset on July 22, 2022

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 the first five chapters of Sakurai's "Modern Quantum Mechanics", as well as some additional topics. If any errors are found in the notes, feel free to email me at ryoweil6@student.ubc.ca.

Contents

1	Fundamental Concepts	2
1.1	The beginnings of quantum mechanics	2
1.2	Kets, Bras, and Hilbert Space	6
2	Quantum Dynamics	11
3	Quantum Foundations	12
4	Symmetries and Angular Momentum	13
5	Approximation Methods	14
6	Identical Particles	15

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):

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

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):

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

which agreed with high wavelength/low frequency data well but failed to accurately describe low wavelength/high frequency emission¹. 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:

$$E = hf. \quad (1.3)$$

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

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

which agrees with the BBR spectrum data across all frequencies². 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 value³:

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

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 out⁴.

¹It 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.

²Further, 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.

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

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

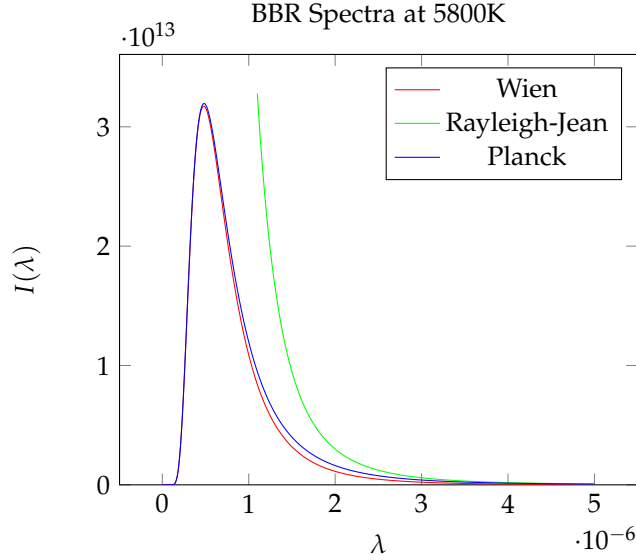


Figure 1.1: Plots of the black body emission spectra at $T = 5800\text{K}$ (the approximate temperature of the surface of the sun) as predicted by Wien's Law, Rayleigh-Jean's Law, and Planck's Law. Planck's Law was found to agree with experimental observations for all wavelengths. Wien's Law agrees with 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:

$$p = \frac{h}{\lambda}. \quad (1.6)$$

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:

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

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 inhomogeneous 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-1/2⁵. This provides the silver atom with a net magnetic moment μ proportional to the electron spin⁶ \mathbf{S} :

$$\mu \propto \mathbf{S}. \quad (1.8)$$

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

$$E = -\mu \cdot \mathbf{B}. \quad (1.9)$$

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

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

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:

$$F_z = \mu_z \frac{\partial B_z}{\partial z}. \quad (1.11)$$

So in the inhomogeneous 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 μ can point in any direction, and therefore μ_z ranges continuously from $+|\mu|$ to $-|\mu|$. 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

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

⁶The 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 negligible 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.

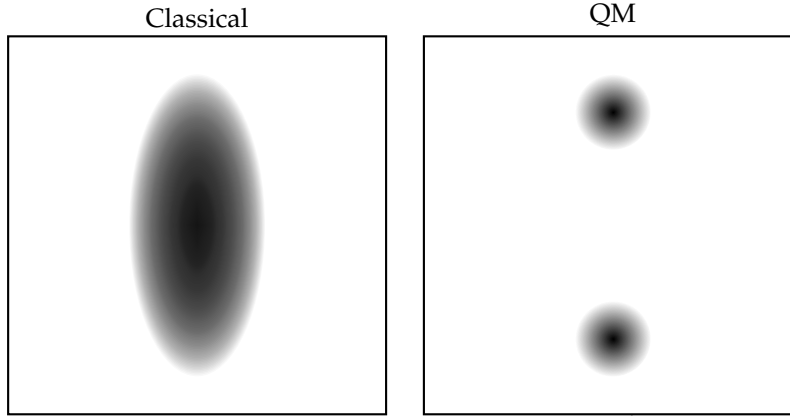


Figure 1.4: Classical prediction (left) and quantum mechanical prediction (right) for the Stern-Gerlach experiment. The screen on the right was observed in experiment.

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.

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle \quad (1.12)$$

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:

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

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 $\pm x$:

$$|x, \pm\rangle = \frac{|\uparrow\rangle \pm |\downarrow\rangle}{\sqrt{2}}. \quad (1.14)$$

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 correspondence⁸ between the kets $|\psi\rangle \in \mathcal{H}$ and the bras⁹ $\langle\psi| \in \mathcal{H}^*$. We call this the *dual correspondence*, and write it as follows:

$$|\psi\rangle \overset{DC}{\longleftrightarrow} \langle\psi|. \quad (1.15)$$

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:

⁷This 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.

⁹Formally, 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.

⁹Given 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:

$$\langle\varphi|\psi\rangle \in \mathbb{C} \quad (1.16)$$

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:

$$0 \leq |\langle\varphi|\psi\rangle| \leq 1. \quad (1.17)$$

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 $\mathcal{B} = \{|b_j\rangle\}_j$ such that every state $|\psi\rangle \in \mathcal{H}$ can be written in the form:

$$|\psi\rangle = \sum_j \psi_j |b_j\rangle \quad (1.18)$$

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.

¹⁰In 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 $\mathcal{B} = \{|b_j\rangle\}_j$ is *orthonormal* if

$$\langle b_i | b_j \rangle = \delta_{ij}. \quad (1.19)$$

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

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

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:

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

We will now prove a useful trick involving ONBs.

Theorem: Resolution of the identity

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

$$\sum_j |b_j\rangle \langle b_j| = \mathbb{I} \quad (1.22)$$

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 $\{|b_j\rangle\}_j$ of \mathcal{H} . Further, recall that $\psi_j = \langle b_j | \psi \rangle$ if the basis is orthonormal. Hence we have that:

$$|\psi\rangle = \sum_j \langle b_j | \psi \rangle |b_j\rangle = \sum_j |b_j\rangle \left(\langle b_j | \psi \rangle \right) = \sum_j \left(|b_j\rangle \langle b_j| \right) |\psi\rangle = \left(\sum_j |b_j\rangle \langle b_j| \right) |\psi\rangle. \quad (1.23)$$

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. \square

The above discussion has been extremely abstract and mathematical so far; let us now reconnect with the physics. In a quantum measurement, the alternatives $\{|m_j\rangle, j \in \text{Outcomes}\}$ 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 $\{|m_j\rangle\}_j$, with:

$$\langle m_i | m_j \rangle = \delta_{ij}. \quad (1.24)$$

Note that $\{|m_j\rangle\}_j$ 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:

$$p_j = |\langle m_j | \psi \rangle|^2. \quad (1.25)$$

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 above statement requires some refinement to deal with cases of degeneracy, which we will soon address. The second question concerns the post-measurement states $\{|m_j\rangle\}_j$; namely, how are they 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 define linear operators, which describe quantum mechanical observables and evolution.

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:

$$A(\alpha|\psi\rangle + \beta|\varphi\rangle) = \alpha(A|\psi\rangle) + \beta(A|\varphi\rangle) \quad (1.26)$$

$\forall |\psi\rangle, |\varphi\rangle \in \mathcal{H}$ 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:

$$(A + B) + C = A + (B + C) \quad (1.27)$$

$$(AB)C = A(BC) \quad (1.28)$$

$$A(B + C) = AB + AC. \quad (1.29)$$

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

$$AB \neq BA. \quad (1.30)$$

2 Quantum Dynamics

3 Quantum Foundations

4 Symmetries and Angular Momentum

5 **Approximation Methods**

6 Identical Particles