A Mathematical Development of Quantum Mechanics

A Thesis in Mathematics

Michael Clancy

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Bachelor in Arts
With Specialized Honors in Mathematics

May 2018

ABSTRACT

The mathematics necessary for an axiomatic statement of the laws of quantum mechanics are introduced and developed. The axioms of quantum mechanical systems are stated. A novel derivation of the projection valued measures associated with the position and momentum observables, as well as the Hamiltonian operator for the quantum harmonic oscillator are provided. Using these projection valued measures, the familiar laws and probabilities of quantum mechanics are shown to be equivalent to the abstract formulation in the axioms.

CONTENTS

1. INTRODUCTION

While sitting in the lectures of my undergraduate quantum mechanics course, I could never shake this profoundly uneasy feeling. It wasn't, as one might expect, the strange physical ramifications of the theory, e.g. uncertainties and entanglement. Rather, it was the non-rigorous and ambiguous statement of the theory itself that bothered me. The level at which I was learning the theory was not self-sufficient: propositions felt unjustified, definitions incomplete, and my bearing on problems felt aimless. I finished the course unsatisfied, even believing that the theorists themselves still had work to do on proving some of their own propositions.

My confusion with quantum mechanics started when I was twelve years old. I read once that scientists believed the universe to be random. I figured they just saw that they couldn't predict something, from which they must have concluded that there was simply no method for predicting anything. I knew very little of the intricacies of the theory of quantum mechanics, and I already thought myself qualified to shut down decades of geniuses based on ten minutes of reading. As it would turn out, the theory is almost predictive, but its nuances are subtle and can only be understood once one has a complete grasp of the necessary mathematics.

This paper sets out to address this issue as concisely as possible. In order to even write down the equations, we must build up the definitions for the objects we put on both sides of each equation. This includes all of the usual suspects: Hilbert spaces, self-adjoint operators[1], spectra, and unitary operators. This innocuously short list sits atop a massive scaffolding of theorems and definitions. For example, one might ask: is the operator $Q\psi = x\psi$ self-adjoint? Well, as it turns out, this is a malformed question. One must first understand the domain upon which an operator may act before asking about any of its properties. It is, in fact, even possible for an operator to not have a well-defined adjoint. This issue, as well as numerous others, stems from the trouble of infinities: general quantum mechanical problems are associated with infinite dimensional Hilbert spaces. Linear operators on infinite dimensional Hilbert spaces can't always be defined on the entire Hilbert space, as they can with finite dimensional Hilbert spaces. In fact, the position operator and momentum operator aren't even defined on the same dense domain, so when one writes down the famous commutation relation [1]

[1] Usually, the term "Hermitian" is used to describe the operators in quantum mechanics. We avoid this term, so as to avoid ambiguity, as the way it is usually used looks more like our definition for "symmetric", which will be given in a later section.

$$[Q, P] = i\hbar \tag{1.1}$$

for position and momentum operators Q and P, and reduced Planck's constant $\hbar$, one must be very careful to define what is *even meant* by $[,]$ for unbounded operators. Similarly, when ones solves the time independent Schrödinger equation of a system, i.e. when given some potential function V, for particle of mass m, one solves the equation

$$\frac{-\hbar^2}{2m}\psi'' + V\psi = E\psi \tag{1.2}$$

for "energy eigenvalues" E and "functions" ψ, it will often be the case that ψ doesn't lie in the Hilbert space, or worse: sometimes ψ isn't even a function at all! Quantum mechanics is very often done this way. Case in point: Take the position operator. It is often stated that the eigenvalues of the position operator are the whole real line, with the eigenvalue x_0 corresponding to the eigenvector $\psi(x) = \delta(x - x_0)$. This is simply not true by the standard definition of eigenvectors and eigenvalues for Hilbert spaces. What is true is that the *spectrum* of the position operator is the whole real line. The *spectrum* of an operator in infinite dimensions is a generalization of the notion of the set of eigenvalues of an operator in finite dimensions, and reduces to the set of eigenvalues in the finite dimensional case.

What is typically done, however, is speak rigorously of finite dimensional problems and posit that most of the tricks work (with possibly some slight modifications) in the infinite dimensional case. For example, for certain operators, like the harmonic oscillator Hamiltonian H, one may construct an orthonormal basis of functions f_n using just the eigenfunctions of H. Orthonormality here means:

$$\langle f_n, f_m \rangle = 1 \quad \text{if} \quad n = m \quad \text{and} \quad 0 \quad \text{otherwise}$$

When one wishes to discuss the position operator, it is often stated that the eigenkets [1] of the position operator are orthogonal in a different way:

$$\langle x|x' \rangle = \delta(x - x') = \infty \quad \text{if} \quad x = x' \quad \text{and} \quad 0 \quad \text{otherwise}$$

This definition turns out to be very useful for certain calculations, but now a curious eye leads one to several questions: are these delta functions Hilbert space elements? (Answer: No.) Does every operator have an orthonormal basis? Can it always be checked for whether the number of elements in the basis is countable or uncountable? Can an operator have a basis which is some parts countable, and some parts uncountable? As it turns out, Dirac developed a very tidy formalism for dealing with questions like this [2]. However, when trying to prove most of his assertions, one is left wanting.

Our version of the operator decomposition is the spectral theorem, an indispensable tool for stating most of the axioms of quantum mechanics. Equipped

with the spectral theorem, we may confidently address every issue we may personally have or a critic may have for our theory, and go beyond simply saying "it works": we will know why it works, where it fails, and what the limitations of our own language are.

In this vein, wouldn't it be nice if there were several axioms which we could call "quantum mechanics"? If there were a way could eliminate all ambiguity and confusion and anchor our understanding to these several statements? As it turns out, yes. We will start by stating these axioms, and explore them in each proceeding section. This paper serves as a complete introduction to the mathematics of quantum mechanics, and applies this rigorous formalism to understand the measurement and time evolution of the quantum harmonic oscillator for particles in $\mathbb{R}$. The framework developed may be applied to any system, but this serves as an illustrative and practical examples of the precise statement of quantum mechanics.

We do not wish to motivate physically, or even logically, the rules of quantum mechanics as they appear in the formalism. Why should the rules be derivable from intuition? For example, one might say that we use Hilbert space to deal with superpositions of states, e.g. energy eigenstates. However, the energy states correspond to wave-functions, which are Hilbert space elements in the first place. One could equivalently say that Hilbert space being the setting for states does not care so much about the observer finding the system in a definite energy eigenstate, or speaking of general states as superpositions of them. The states of the system are independent of the basis chosen by the observer. The latter perspective is the perspective we take: The rules, however they may be thought about, derived, or formulated, are simply the rules. We study them as they are mathematically, so that we may understand them unambiguously.

In each chapter, we will delve briefly into several areas of mathematics: e.g. Hilbert spaces, analysis, and measure theory.

The second chapter of this paper will state the axioms of quantum mechanics. The third chapter of this paper will focus on developing the central player of quantum mechanics: Hilbert spaces. Every quantum mechanical system is associated with a certain Hilbert space. States, observables, time evolution, and measurement statistics of a system are all associated with operators acting on the Hilbert space. The fourth chapter will thus focus on operators on a Hilbert space and their spectra. The fifth chapter will provide a brief introduction to measure theory necessary for the definition of Lebesgue integration and for the development of the spectral theorem. The sixth chapter will introduce several types of functions and spaces of functions necessary for quantum mechanics. The seventh chapter will introduce the notion of projection valued measures and state the spectral theorem. The eighth chapter will introduce observables formally, and the ninth chapter will apply the machinery developed to the quantum harmonic oscillator.

This work is largely based on a lecture series by Frederic P. Schuller [5]. Most of the content is unchanged, but I have presented the material in a completely different order, with more detail. A number of proofs of the propositions are new, and I will call attention to these proofs. In particular, I have derived the

form of projection valued measures of the position and momentum operators, as well as the projection valued measure for the Hamiltonian of the quantum harmonic oscillator.

2. THE AXIOMS

The axioms of quantum mechanics can be stated in the following way, and are due to Schuller [5]. Every question about quantum mechanics may be answered, directly or indirectly, according to the following axioms.

Axiom 1. *Every quantum mechanical system is associated with a complex Hilbert space* $\mathcal{H}$*.*

Axiom 2. *The states of a quantum system* H *are represented by trace class positive linear operators* $\rho : \mathcal{H} \longrightarrow \mathcal{H}$ *such that*

$$Trace\, \rho = 1 \tag{2.1}$$

Axiom 3. *Every observable quantity of a system is represented by a self-adjoint operator* $A : D_A \longrightarrow \mathcal{H}$*.*

Axiom 4. *Given a quantum mechanical system* $\mathcal{H}$ *in the state* ρ*, the probability of obtaining a measurement value in the set* $E \subseteq \mathbb{R}$ *for the observable* $A : D_A \longrightarrow \mathcal{H}$ *is given by*

$$Prob\,(E) = Trace\,(\rho P_A(E)) \tag{2.2}$$

where P_A *is the unique projection valued measure such that*

$$A = \int \lambda dP_A \tag{2.3}$$

Axiom 5. *(Unitary Dynamics) Given a quantum system* $\mathcal{H}$ *in the state* ρ *at time* t_0*, the state at any time* t *is given by*

$$\rho(t) = U(t - t_0)\rho(t_0)U(t - t_0)^{-1} \tag{2.4}$$

where $U(t) = e^{\frac{-iHt}{\hbar}}$ *(defined according to the spectral theorem) and* $H : D_H \longrightarrow \mathcal{H}$ *is the Hamiltonian observable of the system.*

Axiom 6. *(Projective Dynamics) Given a quantum system* $\mathcal{H}$ *in the state* ρ *at time* t_0*, the state of the system after performing a measurement of the observable* $A : D_A \longrightarrow H$ *is given by*

$$\rho(t_0 + \epsilon) = \frac{P_A(E)\rho(t_0)P_A(E)}{Trace(numerator)} \tag{2.5}$$

where E *is the smallest measurable set in which the value of the observable necessarily existed.*

Now, there is a lot to be said about these axioms, and all of the definitions necessary for their precise statement. In fact, the rest of this paper is devoted to defining and exploring all of these axioms. However, I will make a few comments before we really get into it.

First, sometimes I use language that would suggest the physical things *are* their mathematical representation. Within the context of these axioms, we assert this identification. However, the world in which we live is relativistic, and this list of axioms does not take relativity into account. That is to say, if there were a complete list of mathematical truths about reality, this isn't it. It is, however, a very rigid mathematical framework that, not only proves to be almost always true (most systems we deal with have pretty low speed and mass), but locks into place what we believe about the world in the absence of relativity and quantum field theory, so that if those two ever achieve mathematical unification and a precise mathematical formulation, we can unambiguously differentiate them from our six axioms here.

In order to *do* quantum mechanics, one does not need all of this. There are plenty of texts [1, 3, 4] that teach you how to solve all kinds of quantum mechanical problems, and, at the end of the day, carrying around all of this technical formalism gets a little exhausting. You will get the answer much faster if you stick to the formalism of these standard texts during your day job. On Sunday evenings with your colleagues, however, you can shut down any questions they have about the foundations of quantum mechanics by listing them these axioms.

Now that we have stated the axioms, we start from the ground up to build up to them.

3. HILBERT SPACE

Hilbert space plays a central role in quantum mechanics. Just as an understanding of Maxwell's equations requires a thorough knowledge of vectors, a proper treatment of quantum mechanics requires a thorough understanding of Hilbert spaces, which this section aims to provide. In a sense, every quantum mechanical problem is associated with some Hilbert space. Spin systems are associated with the Hilbert space $\mathbb{C}^2$. Particles in space are associated with $L^2(\mathbb{R}^3)$, (sometimes referred to as the "square-integrable functions"). The state of a system, the observable quantities of a system, and even the time evolution of a system are represented by operators acting on the Hilbert space of the system. Hilbert space is an incredibly rich and beautiful mathematical structure, and a rigorous treatment of these quantum mechanical problems requires a very careful look at that structure. Over the course of our discussion, definitions and theorems will be stated exactly. Many proofs will be provided to give better familiarity with these objects, and tricky and more technical proofs will be outsourced to various references.

The definition of a Hilbert space is rather simple once the prerequisite terms used in the definition are understood. It would not make much sense if we were able to write something too detailed down without doing some heavy lifting first. We start with the general notion of a vector space, and build our way up.

Definition 1. *A vector space over a (scalar) field $\mathbb{K}$ is a tuple $(V,+,\cdot)$, where:*

(a) V is a set

(b) $+ : V \times V \longrightarrow V$ is a binary addition operation

(c) $\cdot : \mathbb{K} \times V \longrightarrow V$ is a scalar multiplication operation

and the tuple satisfies the following axioms:

(i) Commutativity: $v + w = w + v$

(ii) Associativity of addition: $(v + w) + u = v + (w + u)$

(iii) Neutral element: $\exists 0 \in V : v + 0 = v$

(iv) Inverse: $\exists - v \in V : v + (-v) = 0$

(v) Associativity of multiplication: $c(dv) = (cd)v$

(vi) Distributivity of scalar sum: $(c + d)v = cv + dv$

(vii) Distributivity of vector sum: $c(v + w) = cv + cw$

(viii) Unity: $1v = v$

and $v, w, u \in V, c, d \in \mathbb{K}$ are all understood to be in the "for all" sense. A "field" is just a set with addition, subtraction, multiplication, and division; the precise definition does not concern us. A helpful mnemonic for the vector space axioms is the meaningless phrase "CANI ADDU". In quantum mechanics, we are interested in vector spaces over $\mathbb{C}$.

We will refer to the vector space tuple, in the following sections, simply by the set "V", when the operations are unambiguously understood; it simply becomes cumbersome to carry around so many symbols.

Proposition 1. $\mathbb{C}^n := \{(z_1, ..., z_n) | z_i \in \mathbb{C}\}$ *is a vector space over* $\mathbb{C}$*, when equipped with the addition operation* $z + w := (z_1 + w_1, ..., z_n + w_n)$*, and the scalar multiplication operation* $c \cdot z = (cz_1, ..., cz_n)$*. These are referred to as "point-wise" addition and multiplication.*

Proof. While proving that something is a vector space, one must first prove that the candidate operations land you back in the vector space. In the following section we enumerate the axioms that require proof as in Definition 1.
(b) Let $z, w \in \mathbb{C}^n$. Then $z + w = (z_1 + w_1, ..., z_n + w_n)$. Since complex numbers are closed under addition, each $z_i + w_i \in \mathbb{C}$. Therefore $z + w \in \mathbb{C}^n$.
(c) Similar to (b); left to the reader.
(i) Commutativity of vector addition follows from commutativity of addition of complex numbers.
(ii - viii) Left to the reader. Hints: (iii) 0 := (0,...,0), (iv) Given z, what is a candidate for $-z$? □

Most of the proofs are trivial in the case of $\mathbb{C}^n$. This is no coincidence; the abstract vector space definition is based on the structure of $\mathbb{C}^n$ and $\mathbb{R}^n$.

Definition 2. *A subset* W *of a vector space* V*, i.e.* $W \subseteq V$ *is said to be a subspace if* W *is also a vector space.*

Proposition 2. *In order to show a subset* W *of* V *is a subspace, it suffices to show:*

(i) Closure under multiplication: $\lambda w \in W$ *for* $w \in W$, $\lambda \in K$

(ii) Closure under addition: $w + u \in W$ *for* $w, u \in W$

(iii) $0 \in W$ *where* 0 *is the* 0 *of* V*.*

Proof. All of the other structure is inherited from V, with the exception of (i) here providing the additive inverse. □

A certain kind of subspace arises very naturally in the following way. Suppose we have a vector space V, and a collection of vectors $v_1, ..., v_n \in V$. By the vector space axioms, $c_1v_1 + ... + c_nv_n$ is still a vector. This kind of vector is given the following name:

Definition 3. *A linear combination of a collection of vectors $v_1, ..., v_n$ is any $c_1v_1 + ... + c_nv_n$ with each $c_i \in \mathbb{C}$.*

Now, we consider all possible linear combinations of those vectors:

Definition 4. *The span of a set of vectors S is the set of all finite linear combinations of vectors in S. That is,*

$$span\, S := \Big\{ \sum_i^n c_i v_i \,\Big|\, c_i \in \mathbb{C}, v_1, ..., v_n \in S. \Big\}$$

Remark 1. *In the case that S is a finite set, observe that*

$$span\{v_1, ..., v_n\} := \{c_1v_1 + ... + c_nv_n \,|\, c_i \in \mathbb{C}\} \tag{3.1}$$

Proposition 3. *The span of a collection of vectors $v_1, ..., v_n \in V$ is a subspace of V.*

Proof. We go through the checklist in Proposition 2.

(i) Let $v \in \operatorname{span} S$. Then $v = c_1v_1 + ... + c_nv_n$. So $\lambda v = (\lambda c_1)v_1 + ... + (\lambda c_n)v_n \in \operatorname{span} S$.

(ii) Let $v, w \in \operatorname{span} S$. Then $v = c_1v_1 + ... + c_nv_n$ and $w = d_1v_1 + ... + d_nv_n$. So $v + w = (c_1 + d_1)v_1 + ... + (c_n + d_n)v_n$.

(iii) $0 = 0v_1 + ... + 0v_n \in \operatorname{span} S$.

□

If the $span\, S = W$, we sometimes say S *spans* W.
Our notation for the span of a (possibly infinite) set is a little sloppy, but the intuition is this. Consider $S = a_1, a_2, ..., a_17$. To get something in the span of S, just take any number of elements of S, like a_1, a_2, a_4, a_12 and form a linear combination of those vectors. If you consider every combination of letters and every possible choice of coefficients, that is the span. The reason we restrict ourselves to finite subcollections is because in the case of infinite sets, we don't want to form infinite sums: we will run into a plethora of convergence issues.

Definition 5. *A vector space V is said to be finite dimensional if there exists a finite set S that spans V. If V is not finite dimensional, it is instead called infinite dimensional.*

Consider the span of two vectors $(1,0,0)$ and $(0,1,0)$ in $\mathbb{R}^3$. This subspace is the xy-plane. The xy-plane is also spanned by the vectors $(1,1,0)$ and $(1,-1,0)$.

Now consider a collection of two vectors, v_1, v_2. If v_1 is just a constant multiple of v_2, then this second vector is redundant in a sense: If I try to probe the vector space by considering linear combinations of these two vectors, I don't get anything new than if I had just started with one of them. In the case of three vectors v_1, v_2, v_3, if v_3 lies in the span of v_1 and v_2, it's another redundant vector. Often we wish to consider vectors which aren't redundant in this sense. We say a collection of vectors $v_1, ..., v_n$ is independent if none of the vectors are in the span of the other vectors. An equivalent condition is taken as the definition:

Definition 6. *A collection of vectors $v_1, ..., v_n$ is said to be linearly independent if the statement $c_1v_1 + ... + c_nv_n = 0$ implies that all the $c_i = 0$.*

Now, if we have some vector space V, and we take an independent set of vectors that span V, we have a very useful way of representing arbitrary vectors by means of a basis:

Definition 7. *A basis $\mathbb{B} = \{e_1, e_2, ..., e_n\}$ for a finite-dimensional vector space V is a linearly independent set that spans V.*

Remark 2. *If $B = \{e_1, ..., e_n\}$ is a basis for V, then every vector $v \in V$ can be written as $v = c_1e_1 + c_2e_2 + ... + c_ne_n$.*

The coefficients are a bit like coordinates, as they give us a way to speak of abstract vectors more concretely.

One extremely important kind of function of vectors is an inner product. The inner product provides a way to compare vectors:

Definition 8. *An inner product (on a complex vector space) is a function $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{C}$ satisfying five axioms:*

(i) $\langle f, f \rangle \in \mathbb{R}$ *and* ≥ 0

(ii) $\langle f, f \rangle = 0$ *iff* $f = 0$

(iii) $\langle f, g \rangle = \overline{\langle g, f \rangle}$

(iv) $\langle f, \lambda g \rangle = \lambda \langle f, g \rangle$ *and* $\langle \lambda f, g \rangle = \bar{\lambda} \langle f, g \rangle$

(v) $\langle f, g + h \rangle = \langle f, g \rangle + \langle f, h \rangle$ *and* $\langle f + h, g \rangle = \langle f, g \rangle + \langle h, g \rangle$

for $f, g, h \in V$ and $\lambda \in \mathbb{C}$.

Another very important function is that of one vector, which defines a notion of "length" for the vector.

Definition 9. *A norm is a function $\| \cdot \| : V \longrightarrow \mathbb{R}$ satisfying four axioms:*

(i) $\|f\| \geq 0$

(ii) $\|f\| = 0 \Rightarrow f = 0$

(iii) $\|\lambda f\| = |\lambda|\|f\|$

(iv) $\|f + g\| \leq \|f\| + \|g\|$ *(Triangle Inequality)*

for $f, g \in V$ and $\lambda \in \mathbb{C}$.
The norm and inner product are, in fact, closely related:

Proposition 4. $\|f\| := \sqrt{\langle f, f\rangle}$ *is a norm. That is, inner products generate norms.*

Proof. Let V be an inner product space. Define the norm as above. This is well defined because $\langle f, f\rangle$ is non-negative. We prove the norm axioms one by one:

(i) follows trivially.
(ii) Suppose $\|f\| = 0$. Then $\langle f, f\rangle = 0$. But by (ii) of the inner product axioms, f must be 0.
(iii) Left to the reader.
Before continuing to (iv), we shall prove the **Cauchy-Schwarz inequality**:

$$\forall f, g \in V, \quad |\langle f, g\rangle| \leq \|f\|\|g\| \tag{3.2}$$

To prove this, let $f, g \in V$. Suppose further that $g \neq 0$: the proof is trivial for $g = 0$. Then $\forall \lambda \in \mathbb{C}$,

$$\|f - \lambda g\|^2 \geq 0 \tag{3.3}$$

. Let us expand the left hand side:

$$\|f - \lambda g\|^2 = \|f\|^2 - \lambda\langle g, f\rangle - \bar{\lambda}\langle f, g\rangle + \lambda\bar{\lambda}\|g\|^2.$$

Now, we proceed carefully: For any choice of f and g, the above statement is true $\forall \lambda \in \mathbb{C}$. Therefore, it is true for the very clever choice of $\lambda = \frac{\langle f,g\rangle}{\|g\|^2}$, which is a complex number because $g \neq 0$ by assumption. Furthermore, by (i) and (ii), $\|g\| > 0$. With this choice of λ, the inequality 3.3 becomes:

$$\|f\|^2 - \frac{|\langle f,g\rangle|^2}{\|g\|^2} \geq 0,$$

from which the Cauchy-Schwartz inequality follows. We are now prepared for (iv).
(iv) Let $f, g \in V$. Then

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2 + 2Re\langle f, g\rangle.$$

But for any complex number z, $Re\, z \leq |z|$, so the above equation becomes (with help of the Cauchy-Schwartz inequality):

$$\|f + g\|^2 \leq \|f\|^2 + \|g\|^2 + 2\|f\|\|g\| = (\|f\| + \|g\|)^2,$$

from which (iv) follows. □

Definition 10. *An inner product space is a tuple of a vector space with an inner product, i.e.* $(V, +, \cdot, \langle \cdot, \cdot \rangle)$.

Proposition 5. $\mathbb{C}^n$ *is an inner product space, when equipped with the inner product*

$$\langle z, w \rangle = \sum_{i=1}^{n} \overline{z_i} w_i \tag{3.4}$$

Definition 11. *A normed vector space is a tuple of a vector space with a norm, i.e.* $(V, +, \cdot, \|\cdot\|)$.

Corollary 1. *Inner product spaces are normed vector spaces with respect to the norm generated by the inner product.*

We are at the gates of Hilbert space. Before proceeding, we make careful note of our progress so far. We started by considering vector spaces over $\mathbb{C}$ in their complete generality. This definition is too broad to make an immediate analogy with $\mathbb{R}^n$; some vector spaces don't have norms, and the norm provides us with a way to define the distance between two vectors: in the case[1] of $\mathbb{R}^3$, the number $\|x - y\|$ is exactly the length of a string extended from the point x to the point y. As we will see in the future, it will be very useful to require that sequences of vectors which get arbitrarily close together (defined by the norm) actually converge to an element of the space. To make this precise, we make the following definitions[2]

Definition 12. *A sequence* f_n *in a normed vector space* V *is said to converge to* $f \in V$ *if:* $\forall \epsilon > 0, \exists N \in \mathbb{N}$*:*

$$\|f_n - f\| < \epsilon \text{ whenever } n > N$$

In this case, we write $f = \lim f_n$.

This is a very natural definition for convergence of an infinite sequence: if you pick some positive number like $\epsilon = 0.1$, then after a while every element of the sequence will be within 0.1 of f. But it's true for every positive number, so the sequence gets and stays as close as you want to the limit f.

A slightly weaker condition for a sequence to "converge" is for it to be a *Cauchy* sequence:

Definition 13. *A sequence* f_n *in a normed vector space* V *is said to be Cauchy if:* $\forall \epsilon > 0, \exists N \in \mathbb{N}$*:*

$$\|f_n - f_m\| < \epsilon \text{ whenever } n, m > N$$

[1] The definition for real vector spaces is to replace $\mathbb{C}$ by $\mathbb{R}$ everywhere above and drop every instance of a conjugation. The curious reader may verify that $\mathbb{R}^3$ is a normed vector space.

[2] These definitions are usually given in terms of metric spaces. However, all normed vector spaces are metric spaces with respect to the metric generated by their norms, and the proof is immediate. As we are not interested in general metric spaces, we restrict our notions of convergence etc. to normed vector spaces.

Definition 14. *A normed vector space V is said to be complete if every Cauchy sequence converges to an element of the space.*

In a sense, complete vector spaces are "filled out". Take, for example, the rationals and the reals. There is a sequence of rationals that get closer and closer to the real number $\sqrt{2}$, but the number $\sqrt{2}$ is not rational, so the rational numbers are not complete. The real numbers, however, are complete, which follows from the construction of Dedekind cuts [7]. We will assume the completeness of the reals for the following proposition:

Proposition 6. $\mathbb{C}$ *is complete.*

Proof. Let z_n be a Cauchy sequence. Then $z_n = x_n + iy_n$, and $\forall \epsilon > 0, \exists N$:

$$|z_n - z_m| < \epsilon \text{ whenever } n, m > N$$

But both $|x_n - x_m|$ and $|y_n - y_m|$ are less than $|z_n - z_m|$, so we have two real Cauchy sequences x_n and y_n. But since $\mathbb{R}$ is complete, $\exists x, y \in \mathbb{R}$ such that $x = \lim x_n$ and $y = \lim y_n$. Define $z = x + iy$. Then:

$$|z_n - z| \leq |x_n - x| + |y_n - y|$$

But both expressions on the right hand side can be made less than $\epsilon/2$ for all n some N_x and N_y, respectively. We therefore have

$$|z_n - z| < \epsilon \text{ whenever } n, m > N := \max(N_x, N_y)$$

Therefore $\lim z_n = z \in \mathbb{C}$ and so z_n converges in $\mathbb{C}$. But this true for any Cauchy sequence, so $\mathbb{C}$ is complete. □

Proposition 7. $\mathbb{C}^n$ *and* $\mathbb{R}^n$ *are complete normed vector spaces.*

Proof. The proofs use the same triangle inequality trick as for $\mathbb{C}$. The details are left to the reader. □

Complete normed vector spaces turn up so often that they are named after someone:

Definition 15. *A Banach space V is a complete normed vector space.*

When the norm is generated by an inner product, we end up with an extremely important object:

Definition 16. *A Hilbert space H a complete inner product space.*

Remark 3. *Hilbert spaces are Banach spaces with norms generated by their inner product.*

That's it. No fuss; no ambiguity: a Hilbert space is any vector space that has an inner product, and the norm generated by the inner product space has the nice property that Cauchy sequences converge. Now that we know what a Hilbert space is, we can construct concrete examples of one.

Corollary 2. *$\mathbb{C}^n$ is a Hilbert space.*

Now, so far, every vector space that we've considered was already a Hilbert space. Are there vector spaces which aren't normed, have no inner product, or aren't complete? If we're talking about finite dimensional vector spaces over a field which is complete, it is not hard to show that we can't construct such vector spaces. It is only when me move to infinite dimensions that we see being a vector space does not guarantee an inner product or completeness (even over a complete field like $\mathbb{C}$). Let us attempt, then, to construct an infinite dimensional Hilbert space from scratch. Let us try to generalize our definition of $\mathbb{C}^n$ to the set of infinite complex sequences:

$$\mathbb{C}^\infty := \Big\{(z_1, z_2, ...) \Big| \, z_i \in \mathbb{C}\Big\}. \tag{3.5}$$

When equipped with the addition operation $(z_1, z_2, ...) + (w_1, w_2, ...) = (z_1 + w_1, z_2 + w_2, ...)$, and the scalar multiplication operation $c(z_1, z_2, ...) = (cz_1, cz_2, ...)$, the reader may verify that $\mathbb{C}^\infty$ is a vector space according to Definition 1. Let's see if $\mathbb{C}^\infty$ is a Hilbert space with a generalization of the usual inner product. We'll try to define:

$$\langle z, w\rangle := \sum_{n=1}^{\infty} \overline{z_i} w_i \tag{3.6}$$

If we try to define the norm with this candidate inner product, we see that a harmless vector like $\xi = (1, 1, ...)$ has $\|\xi\| = 1 + 1 + ... = \infty$, and we're already in trouble: ξ is infinitely far away from the origin!

So let's try to pick a subset of $\mathbb{C}^\infty$ of vectors that are not like ξ: the so-called *square-summable* complex sequences, whose candidate norm does not blow up. Amazingly, this naïve guess hits the target exactly: This set is a Hilbert space.

Theorem 1. *The set of square summable complex sequences*

$$\ell^2 := \Big\{x \in \mathbb{C}^\infty \Big| \sum_{n=1}^{\infty} |x_i|^2 < \infty\Big\} \tag{3.7}$$

is a Hilbert space, when equipped with the inner product in Equation 3.6.

Proof. We begin the proof by showing that ℓ^2 is a subspace of $\mathbb{C}^\infty$. $0 = (0, 0, ...)$ is trivially square-summable, so it remains to show that (i) and (ii) of Proposition 2 are true.
(i) Let $x \in \ell^2$. We wish to prove that λx is also in ℓ^2. We inspect the sum

$$\sum_{i=1}^{\infty} |\lambda x_i|^2 = \sum_{i=1}^{\infty} |\lambda|^2 |x_i|^2 = \lim_{n\to\infty} \sum_{i=1}^{n} |\lambda|^2 |x_i|^2 = \lim |\lambda|^2 S_n \tag{3.8}$$

But $\lim S_n$ converges to $S = \sum |x_i|^2$. Therefore all we have to show is that we can pull numbers out of convergent sequences, a fact that will be generally

useful. Let μ be any complex number, and S_n any complex sequence (keeping in mind that we need it for sums here). Then let ϵ be any positive number. Then $\exists N$:

$$|S_n - S| < \frac{\epsilon}{|\mu|} \text{ whenever } n > N$$

Therefore

$$|\mu S_n - \mu S| = |\mu||S_n - S| < |\mu|\frac{\epsilon}{|\mu|} = \epsilon \text{ whenever } n > N$$

But that means:

$$\lim \mu S_n = \mu S = \mu \lim S_n$$

and so complex numbers can be pulled out of convergent sequences, so $x \in \ell^2$.
(ii) Let $x, y \in \ell^2$. Then $(x+y)_i = x_i + y_i$, and we examine each term:

$$|x_i + y_i|^2 = |x_i|^2 + |y_i|^2 + 2Re(\overline{x_i}y_i) \tag{3.9}$$

But if we look at $|x_i - y_i|^2 \geq 0$, we obtain

$$2Re(\overline{x_i}y_i) \leq |x_i|^2 + |y_i|^2. \tag{3.10}$$

It then follows from Equations 3.9 and 3.10 that:

$$|x_i + y_i|^2 \leq 2(|x_i|^2 + |y_i|^2).$$

Therefore, because each term is point-wise $\leq$,

$$\sum_{n=1}^{\infty} |x_i + y_i|^2 \leq \sum_{n=1}^{\infty} 2(|x_i|^2 + |y_i|^2) = \sum_{n=1}^{\infty} 2|x_i|^2 + \sum_{n=1}^{\infty} 2|y_i|^2 < \infty$$

because of (i) and the fact that $x, y \in \ell^2$ themselves. The sum can be broken up because everything is non-negative. Therefore ℓ^2 is a vector space.
It remains to show ℓ^2 has an inner product and that it is complete. It gets ugly.

We define the inner product as in Equation 3.6:

$$\langle x, y \rangle := \sum_{i=1}^{\infty} \overline{x_i} y_i$$

For this to be well defined, we must guarantee the sequence converges. It would be nice to prove this directly, but as we will see here and in the future, it is much easier to check the convergence of a sequence in a complete vector space by checking the Cauchy criterion: the convergence follows from completeness. Let us check if the sequence of partial sums is Cauchy (assuming $n \neq m$):

$$|S_n - S_m| = \Big| \sum_{i=1}^{n} \overline{x_i} y_i - \sum_{i=1}^{m} \overline{x_i} y_i \Big| = \Big| \sum_{min(n,m)+1}^{max(n,m)} \overline{x_i} y_i \Big|$$

But by the triangle inequality for complex numbers we have

$$\Big| \sum_{min(n,m)+1}^{max(n,m)} \overline{x_i} y_i \Big| \leq \sum_{min(n,m)+1}^{max(n,m)} |\overline{x_i} y_i| = \sum_{min(n,m)+1}^{max(n,m)} |x_i||y_i|$$

How can we make this less than ϵ? Well we inspect the relation

$$(|x_i| - |y_i|)^2 \geq 0$$

from which we can conclude:

$$|x_i||y_i| \leq \frac{1}{2}(|x_i|^2 + |y_i|^2)$$

Therefore, because for each i this is a positive number, the inequality is preserved in the sum:

$$\sum_{min(n,m)+1}^{max(n,m)} |x_i||y_i| \leq \frac{1}{2} \sum_{min(n,m)+1}^{max(n,m)} |x_i|^2 + \frac{1}{2} \sum_{min(n,m)+1}^{max(n,m)} |y_i|^2$$

But because x and y are both square summable, the infinite sums converge, and therefore both of those partial sums are Cauchy.[3] Therefore, both of them can be made less than any ϵ. This means the sequence of partial sums of the inner product is Cauchy, and since $\mathbb{C}$ is complete, the infinite sum converges.

Now that we know the candidate inner product is well defined, we check the inner product criteria.

(i) $\langle x, x \rangle \geq 0$
Every term in the infinite sum is nonnegative for any x, so the whole sum is nonnegative for any x.

(ii) $\langle x, x \rangle = 0 \implies x = 0$
Suppose there is some x that has $\sum_{i=0}^{\infty} |x_i|^2 = 0$. Since every term is non-negative, every term would have to be zero. Therefore the x has to be the zero vector.

(iii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$
This is the same as checking: can the conjugation be pulled inside the infinite sum? Well it can be pulled inside any finite sum, and the infinite sum is the sequence of the finite sums, so this amounts to checking if for any sequence $\lim z_n = z$ we can conclude that $\lim \overline{z_n} = \overline{z}$. Observe that for any complex number w, it is true that $|w| = |\overline{w}|$, and therefore

$$|z_n - z| = |\overline{z_n - z}| = |\overline{z_n} - \overline{z}|$$

So whenever we can make the expression on the left less than ϵ, the one on the right is automatically less than ϵ.

[3] Cauchy is weaker than convergence, so converge always implies Cauchy. Completeness is required to show Cauchy implies convergence

(iv) $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$ and $\langle \lambda x, y\rangle = \overline{\lambda}\langle x, y\rangle$
Constants can be pulled out of convergent infinite sums.

(v) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ and $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$
Infinite sums can be broken up if both sums are convergent.

So far, ℓ^2 is an inner product space. All that remains to show is its completeness. Let x^n be a Cauchy sequence in ℓ^2. Be careful with the notation here: something like x^7 itself is an infinite sequence of complex numbers, so we will denote the elements of that sequence by $x^7_1, x^7_2, x^7_3, ...$ and similarly. Let's stare at the sequence of sequences for a while:

$$x^1_1, x^1_2, x^1_3, x^1_4, x^1_5, x^1_6, ...$$

$$x^2_1, x^2_2, x^2_3, x^2_4, x^2_5, x^2_6, ...$$

$$x^3_1, x^3_2, x^3_3, x^3_4, x^3_5, x^3_6, ...$$

$$...$$

Each row is a square-summable sequence. Row by row, we have a Cauchy sequence. Let's pick just one column. That is, for any L, it's always true that

$$|x^n_L - x^m_L| \leq |x^n - x^m|$$

So we see that for each L the sequence x^n_L is Cauchy in $\mathbb{C}$, and so that sequence converges. We will denote this by $\lim_{n\to\infty} x^n_L = x_L$. Therefore for we define the candidate limit x of the x^n's according to this, by $x_L = \lim_{n\to\infty} x^n_L$. First we must check that this candidate is even in ℓ^2 so we can put it inside the norm and check if it's the limit. To check for its square-summability, we have to be rather clever.
If we try to write something like

$$\sum_{i=0}^{\infty} |x_i|^2 = \sum_{i=0}^{\infty} |\lim_{n\to\infty} x^n_i|^2$$

we worry about all kinds of convergence issues and infinities. Instead, all we have to do is bound x's square sum above. To accomplish this, we use just about the only fact we do know about x: the sequence defining it is Cauchy. Since x^n is Cauchy, then for any ϵ we can find an N that makes $\|x^n - x^m\| < \epsilon$ for $n, m > N$. Now comes the first clever step: we make the choice $\epsilon = 1$.[4] Then $\exists N$:

$$\|x^n - x^m\| < 1 \,\text{whenever}\, n, m > N$$

But squaring both sides preserves inequality, so

$$\|x^n - x^m\|^2 < 1 \,\text{whenever}\, n, m > N$$

[4] we could choose $\epsilon =$ anything, but we will just need a fixed positive real number to bound the sum above.

But any $x^n - x^m$ is still in ℓ^2 because ℓ^2 is a vector space. We know what the norm-squared of a vector looks like:

$$\|x^n - x^m\|^2 = \sum_{i=0}^{\infty} \|x_i^n - x_i^m\|^2 < 1 \,\text{whenever}\, n, m > N$$

Now comes the second clever step: Each term in the infinite sum is non-negative, so for any $\eta \in \mathbb{N}$,

$$\sum_{i=0}^{\eta} \|x_i^n - x_i^m\|^2 \leq \sum_{i=0}^{\infty} \|x_i^n - x_i^m\|^2$$

So for any η, we have

$$\sum_{i=0}^{\eta} \|x_i^n - x_i^m\|^2 < 1 \,\text{whenever}\, n, m > N$$

Now we fix η to be anything and fix $n > N$, and take the limit on m:

$$\lim_{m\to\infty} \sum_{i=0}^{\eta} \|x_i^n - x_i^m\|^2 \leq 1 \,\text{whenever}\, n > N$$

(When taking the limit, inequalities pick up an "or equals".) Since η is fixed, we can pull the limit inside the finite sum with ease, as well as inside the norm:

$$\lim_{m\to\infty} \sum_{i=0}^{\eta} \|x_i^n - x_i^m\|^2 = \sum_{i=0}^{\eta} \|x_i^n - \lim_{m\to\infty} x_i^m\|^2 = \sum_{i=0}^{\eta} \|x_i^n - x_i\|^2 < 1 \,\text{whenever}\, n > N$$

For each i, we have from the reverse triangle inequality that:

$$\|x_i\|^2 - \|x_i^n\|^2 \leq \|x_i^n - x_i\|^2$$

And since each term is point-wise not bigger, we have the same in the sum:

$$\sum_{i=0}^{\eta} \|x_i\|^2 - \|x_i^n\|^2 \leq \sum_{i=0}^{\eta} \|x_i^n - x_i\|^2 < 1 \,\text{whenever}\, n > N$$

Now, we fix $n = N + 1$, rearrange, and take the limit on η:

$$\lim_{\eta\to\infty} \sum_{i=0}^{\eta} \|x_i\|^2 \leq 1 + \lim_{\eta\to\infty} \sum_{i=0}^{\eta} \|x_i^{N+1}\|^2$$

But since x^{N+1} is square-summable, the term on the right is finite and so:

$$\sum_{i=0}^{\infty} \|x_i\|^2 < \infty$$

So $x \in \ell^2$.
To show x is in fact the limit, we use some similar techniques. First let us fix an η. Observe that:

$$\sum_{i=0}^{\eta} \|x_i^n - x_i^m\|^2 \leq \sum_{i=0}^{\infty} \|x_i^n - x_i^m\|^2$$

for any choice of n and m. As well, for any $\epsilon > 0$, we may find an N so that:

$$\sum_{i=0}^{\infty} \|x_i^n - x_i^m\|^2 < \frac{\epsilon^2}{4} \text{ whenever } n, m > N$$

Which guarantees for any η that:

$$\sum_{i=0}^{\eta} \|x_i^n - x_i^m\|^2 < \frac{\epsilon^2}{4} \text{ whenever } n, m > N$$

Now if we take the limit on m we pick up an equals sign:

$$\lim_{m\to\infty} \sum_{i=0}^{\eta} \|x_i^n - x_i^m\|^2 \leq \frac{\epsilon^2}{4} < \frac{\epsilon^2}{2} \text{ whenever } n > N$$

So pull the limit in for any choice of η and any choice of $n > N$. Then

$$\sum_{i=0}^{\eta} \|x_i^n - x_i\|^2 < \frac{\epsilon^2}{2} \text{ whenever } n > N$$

But our choice of N guaranteed that any choice of η satisfies this inequality. So taking the limit on η we pick up an equals sign:

$$\sum_{i=0}^{\infty} \|x_i^n - x_i\|^2 \leq \frac{\epsilon^2}{2} < \epsilon^2 \text{ whenever } n > N$$

The term on the left is just $\|x_n - x\|^2$. Therefore we have $\forall \epsilon > 0$, $\exists N$:

$$\|x_n - x\| < \epsilon \text{ whenever } n > N$$

And so $x = \lim x_n$. Therefore, every Cauchy sequence converges in ℓ^2, and ℓ^2 is complete. So ℓ^2 is a complete inner product space: a Hilbert space. □

3.1 Dense Subsets and Topology

Very frequently in dealing with Hilbert spaces, it will be impossible to define linear operators on the entire Hilbert space. Instead, it will suffice to define it on a very special kind of subset of the Hilbert space: dense subspaces. A dense subset, loosely speaking, is a subset that has points *basically* everywhere in the Hilbert space. They are useful because many such dense subsets are ones which have extremely nice and workable properties, and often operators can be defined simply on dense subsets, and then *extended* to the rest of the Hilbert space.

Definition 17. *Given a Hilbert space H, a subset $M \subseteq H$ is said to be dense in H if $\forall \psi \in H, \exists \psi_n \in M$ such that $\lim \psi_n = \psi$. That is, every point in the Hilbert space can be written as the limit of a sequence in the subset.*

Some basic topological notions provide an easier way of talking about dense subsets:

Definition 18. *Given a normed vector space V, we define the neighborhood with radius r around a point v_0 by:*

$$N_r(v_0) = \left\{ v \in V \middle| \|v - v_0\| < r \right\}$$

The neighborhoods around a point are surface-less solid spheres with their centers at that point.

Definition 19. *A deleted neighborhood ${N_r}'(v_0)$ around a point v_0 is a neighborhood minus the point. That is,*

$${N_r}'(v_0) = \left\{ v \in V \middle| 0 < \|v - v_0\| < r \right\}$$

Proposition 8. *Let H be a Hilbert space, and let $M \subseteq H$. Then M is dense in H iff $\forall \psi \in H$, it is true that every neighborhood of ψ contains a point from M.*

Proof. Suppose M is dense in H. Then we know $\exists \psi_n \in M$ such that $\lim \psi_n = \psi$. So let $N_\epsilon(\psi)$ be a neighborhood around ψ. Then clearly $\exists N$ such that $\|\psi_n - \psi\| < \epsilon$ whenever $n > N$, so $\psi_{N+1} \in N_\epsilon(\psi)$. The converse is similar and left to the reader. □

Definition 20. *Given a normed vector space V, a set $F \subseteq V$ is said to be open if every point has a neighborhood around it that's still contained in the set. That is, F is open if $\forall v \in F, \exists r > 0$: $N_r(v) \subseteq F$.*

Definition 21. *Given a normed vector space V, A set $G \subseteq V$ is said to be closed if G^c is open.*

Definition 22. *A limit point v_0 of a set E is a point that contains points in E that are arbitrarily close to it. That is, v_0 is a limit point of E if $\forall r > 0, \exists v \in {N_r}'(v_0)$: $v \in E$. The collection of limit points of E is denoted by E'.*

Proposition 9. *If v is a limit point of M, then $\exists v_n \in M$ with $\lim v_n = v$.*

Proof. Let $r_n = 1/n$. Then there has to be a v_n within r_n of v. Pick one for each n, so that $\lim v_n = v$. □

Definition 23. *The closure of a set E is defined as $\overline{E} := E \bigcup E'$.*

Proposition 10. *For any set E, $\overline{E}$ is closed.*

Proof. The proof can be found in [7]. □

Proposition 11. *Given a Hilbert space H, a subset $M \subseteq H$ is dense in H iff $\overline{M} = H$.*

Proof. The proof is rather trivial, since the definition of denseness and closure are not very different. Suppose $\overline{M} = H$. We wish to show any point in H can be written as the limit of a sequence in M. Suppose first that ψ is already inside of M. Define $\psi_n = \psi$, and clearly $\lim \psi_n = \psi$ and $\psi_n \in M$. Now suppose $\psi \notin M$. Then since $M \bigcup M' = H$, $\psi \in M'$. Therefore $\exists \psi_n \in M$ with $\lim \psi_n = \psi$. □

3.2 An Infinite Dimensional Basis

If we were to guess at a basis for ℓ^2, an obvious candidate would be the infinite collection $e_1 = (1, 0, 0, ...), e_2 = (0, 1, 0, ...), e_3 = (0, 0, 1, ...), ...$ However, our definition of a basis only exists in finite dimensions, so we must first generalize our notion of basis to include infinite dimensions. There are a few different ways to do this. We define one way exactly like in finite dimensions:

Definition 24. *A Hamel basis $\mathcal{B}$ for an infinite dimensional vector space V is a linearly independent set with $span(\mathcal{B}) = V$.*

This definition is the exact same as a regular basis for finite dimensional vector spaces. However, this notion is rather weak, as our candidate basis above does not work. We may only form finite sums with a Hamel basis, but we want an infinite sum to take care of infinitely many numbers in the sequence, i.e. any v can be written as $\sum_{n=1}^{\infty} c_i e_i$. However, an infinite sum is a sequence of partial sums, and partial sums are finite so they are in the span. This means that v can be expressed as the limit of a sequence in the span, which is exactly the condition for denseness.

Definition 25. *A Schauder basis (B) for an infinite dimensional Banach space V is a linearly independent set with $\overline{span(\mathcal{B})} = V$. That is, the span of $\mathcal{B}$ is dense in the whole space.*

Definition 26. *A Hilbert space is said to be separable if there exists a countable dense subset. That is, $\exists A$, countable, with $\overline{A} = H$.*

Remark 4. *From now on, when we say "basis", we use the notion of the Schauder basis. The Schauder basis is much more useful for quantum mechanics.*

Remark 5. *After careful inspection of the definition of a Schauder basis for separable Hilbert spaces, one may see that it says if we have a Schauder basis ψ_n, we may write for each $\psi \in H$ that there exists coefficients c_n such that*

$$\psi = \sum_{n=1}^{\infty} c_n \psi_n$$

In a sense, separable Hilbert spaces are not much more than countable sets.

Proposition 12. *If a Hilbert space H is separable, then there exists a countable Schauder basis for the space.*

Proof. Can be found in [9]. □

Proposition 13. ℓ^2 *is separable.*

Proof. find the proof □

3.3 Closed Subspaces

Closed subspaces are of particular interest for operators, and a number of theorems will be useful.

Proposition 14. *Given a separable Hilbert space H, a closed subspace $M \subseteq H$ is also separable.*

Proof. If H is separable, then there exists a dense subset $A \subseteq H$ such that A is countable. Now, define the set $B = M \cap A$. The $B \subseteq A$ so that B must be countable as well. Now, let $\psi \in M$. We split this up into two cases: either ψ is in A or it isn't.

If $\psi \in A$, then $\psi \in B$ so that if we define $\psi_n = \psi$ we have the condition that ψ can be written as the limit of a sequence.

Now, suppose $\psi \notin A$. Then since H is separable $\exists \psi_n \in A$ such that $\psi = \lim \psi_n$. Now, since M is a closed subspace, define $(\psi_n)_\parallel$ in the usual way. Because of the fact that $\|(\psi_n)_\parallel - \psi\| < \|\psi_n - \psi\|$ we have $\lim(\psi_n)_\parallel = \psi$. This completes the proof. □

3.4 Orthogonality

Orthogonality is a generalization of the notion of perpendicularity of ordinary $\mathbb{R}^3$ vectors. The inner product allows us to define it in the obvious way:

Definition 27. *Two vectors v, w in an inner product space are said to be orthogonal, or perpendicular, if $\langle v, w\rangle = 0$.*

Definition 28. *We say a collection of vectors $\{v_1, v_2, , , \}$ is orthonormal (ON) if each $\langle v_i, v_j\rangle = 0$ if $i \neq j$ and 1 if $i = j$.*

Definition 29. *Given a Hilbert space H and a subset $M \subseteq H$, the orthogonal complement of M is defined by:*

$$M^\perp := \left\{ \psi \in H \middle| \langle \psi, \eta\rangle = 0 \forall \eta \in M \right\}$$

That is, the elements of the orthogonal complement of M are the vectors that are perpendicular to everything in M.

Proposition 15. *Let $\mathcal{H}$ be a Hilbert space. Given a subspace $A \subseteq \mathcal{H}$, $A^\perp$ is closed.*

Proof. In order to show $A^\perp$ is closed, it only needs to be shown that $\overline{A^\perp} \subseteq A^\perp$. So, let $b' \in \overline{A^\perp}$. Then $b' = \lim b_n$ for $b_n \in A^\perp$. Therefore for any $a \in A$ we have

$$\langle b', a\rangle = \langle \lim b_n, a\rangle = \lim\langle b_n, a\rangle = \lim 0 = 0$$

so $b' \in A^\perp$. □

Proposition 16. *Given a Hilbert space H, and closed linear subspace M, it is true that*

$$H = M \bigoplus M^\perp$$

where by $\bigoplus$ we mean that every element $\psi \in H$ can be uniquely written as $\phi + \eta$ for $\phi \in M$ and $\eta \in M^\perp$.

Proof. Can be found in [6]. □

Lemma 1. *Let $\mathcal{H}$ be a Hilbert space and let $A \subseteq H$ be a subspace. Then*

$$A^\perp = \overline{A}^\perp$$

Proof. Inclusion from right to left holds trivially. To show inclusion in the other direction, let $b \in A^\perp$, and let $\alpha \in \overline{A}$. Then $\alpha = \lim a_n$ for $a_n \in A$ and

$$\langle b, \alpha\rangle = \langle b, \lim a_n\rangle = \lim\langle b, a_n\rangle = \lim 0 = 0$$

□

Proposition 17. *Let $\mathcal{H}$ be a Hilbert space. Given a subspace $A \subseteq \mathcal{H}$,*

$$A^{\perp\perp} = \overline{A} \tag{3.11}$$

Proof. In order to show two sets are the same, we show inclusion in both directions. Let $\alpha \in A^{\perp\perp}$. Since $H = \overline{A} \bigoplus \overline{A}^\perp = \overline{A} \bigoplus A^\perp$, we know $\alpha = a' + b$ for $a' \in \overline{A}$ and $b \in A^\perp$. But this means $b = \alpha - a'$, so $b \in \overline{A}$. Therefore $b = 0$, and $\alpha = a' \in \overline{A}$. Now, let $a' \in \overline{A}$, and let $b \in A^\perp$. Then

$$\langle a', b\rangle = \langle \lim a_n, b\rangle = \lim\langle a_n, b\rangle = \lim 0 = 0$$

So $a' \in A^{\perp\perp}$. This completes the proof.

□

4. LINEAR OPERATORS

Linear operators are, so to speak, the name of the game for quantum mechanics. Measurable quantities are called observables, and are represented by self-adjoint linear operators. The values that can be measured given by the spectrum of that observable. The time evolution of a system is governed by a family of unitary operators parameterized by a time variable, and is generated by the self-adjoint Hamiltonian operator of the system. This section aims to provide a backdrop for all of these definitions and prove the main results.

Operators are sometimes called maps or transformations, but an operator is just a new label for a function between vector spaces. The important part is the linearity:

Definition 30. *An operator $T : V \longrightarrow W$ is said to be linear if:*

(i) $T(v + u) = T(v) + T(u)$

(ii) $T(kv) = kT(v)$

for $v, u \in V$ and $k \in \mathbb{C}$.

In other words, linear operators are general objects that satisfy the simple axioms that regular scalar multiplication satisfies. In particular, given some vector space V, the operator T defined by $T(v) = kv$ is a linear for any k.

Proposition 18. *If $A : V \longrightarrow W$ is a linear operator, then $A(0) = 0$.*

Proof. Let $v \in V$. Then:

$$A(0) = A(v - v) = Av + A(-v) = A(v) - A(v) = 0$$

□

Proposition 19. *The operator $0 : V \longrightarrow W$ defined by $0v = 0$ is linear.*

Proof. Left to the reader. Hints: $0 + 0 = 0$, $k0 = 0$. □

Remark 6. *We use the same symbol 0 to denote the scalar 0, the vector 0, and the operator 0. When the possibility of ambiguity arises we shall write which 0 we mean, e.g. "the vector 0".*

Frequently, we drop the (), and just write Tv to denote $T(v)$.

Two operators T and S are said to be equal if their domains coincide, i.e. $D_T = D_S$, and $T(v) = S(v)$, $\forall v \in D_T$. Often two operators will be basically equal but one domain will be smaller:

Definition 31. *We say $S : D_S \longrightarrow W$ is an extension of $T : D_T \longrightarrow W$ if:*

(i) $D_T \subseteq D_S$

(ii) $T(v) = S(v)$, $\forall v \in D_T$.

In this case, we write $T \subseteq S$, and also say T extends S.

Remark 7. *Two operators are equal if they are both extensions of each other.*

Proposition 20. *Given vector spaces V and W, the set of linear operators*

$$L(V, W) := \{T : V \longrightarrow W \,|\, T \text{ is linear}\}$$

is a vector space, when equipped with addition operation

$$(S + T)(v) := S(v) + T(v) \tag{4.1}$$

and scalar multiplication operation

$$(kS)(v) := k(S(v)). \tag{4.2}$$

When V = W, we sometimes just write $L(V, W) \equiv L(V)$. Be very careful with the definitions here. In the case of addition, the + on the left hand side is different from + on the right hand side of Equation 4.1. The RHS + adds vectors in the vector space. The LHS + adds operators in the set of operators on that vector space. They are not independent, however: the vector space structure imposed on $L(V, W)$ by the LHS + is inherited from the vector space structure imposed on the vector space W by the RHS +. Similarly for scalar multiplication.

This is a pretty interesting result: the set of linear operators between vectors spaces is again a vector space. Can we make Hilbert spaces of operators? As it will turn out, it will be of great utility to construct a Banach space of operators, and so we try to make a Banach space instead. How can we do that? First, we have to define a candidate norm. There are many ways to do this, but one possibility is this: consider first the norm of a vector in $\mathbb{R}^3$. This is the length of an arrow. Is there a length of an operator A? Well, first let us consider maps from normed vector spaces V into normed vector spaces W, so that we have good footing for defining a norm. Let us consider how much bigger A makes a vector by examining the quantities

$$\frac{\|Av\|}{\|v\|}$$

for any $v \neq 0$. This quantity depends heavily on the vector v. However, if this quantity never gets bigger than some finite number, i.e. if

$$\sup_{v \neq 0} \frac{\|Av\|}{\|v\|} < \infty \tag{4.3}$$

then A never makes v infinitely bigger than itself. In this case we call the operator A bounded.

Definition 32. *A is bounded if Equation 4.3 is true.*

The set of bounded linear operators are nice to work with (especially in the case of limits), and we give them a name.

Definition 33. *The set of bounded linear maps between normed vector spaces is denoted by*

$$\mathcal{L}(V, W) := \left\{ A \in L(V, W) \middle| A \text{ is bounded} \right\}$$

Proposition 21. *$\mathcal{L}(V, W)$ is a subspace of $L(V, W)$.*

Whenever A is bounded, we define its norm by how big it makes vectors at most:

Definition 34. *The operator norm $\| \cdot \| : \mathcal{L}(V, W) \longrightarrow [0, \infty)$ is defined by*

$$\|A\| := \sup_{v \neq 0} \frac{\|Av\|}{\|v\|}$$

for normed vector spaces V and W.

Proposition 22. *Let $A \in \mathcal{L}(V, W)$. Then for any $v \in V$,*

$$\|Av\| \leq \|A\|\|v\| \tag{4.4}$$

Proof. Recall that the operator norm is what the operator does to a vector *at worst*. That is, for any $v \in V$, nonzero:

$$\frac{\|Av\|}{\|v\|} \leq \sup_{v \neq 0} \frac{\|Av\|}{\|v\|} = \|A\|$$

And so $\|Av\| \leq \|A\|\|v\|$ for any $v \neq 0$. But in the case $v = 0$, A is linear so $Av = 0$, and equality holds. Therefore Equation 4.4 holds for all v. □

Proposition 23. *An equivalent definition of the operator norm is to only consider vectors on the unit sphere of the domain. That is,*

$$\|A\| = \sup_{\|v\|=1} \|Av\|$$

Proof. Define

$$\alpha = \left\{ \frac{\|Av\|}{\|v\|} \middle| v \in V, v \neq 0 \right\}$$

and

$$\omega = \left\{ \|Av\| \middle| v \in V, \|v\| = 1 \right\}$$

In order to show $\sup \alpha = \sup \omega$, we show that the two sets are equal. Let $x \in \omega$. Then $x = \|Av\|$ for some v on the unit sphere. Then certainly $v \neq 0$, so

$$\frac{\|Av\|}{\|v\|} = \|Av\| = x \in \omega$$

It remains to show $\alpha \subseteq \omega$. Let $x \in \alpha$. Then there's some $v \in V, v \neq 0$ such that $x = \frac{\|Av\|}{\|v\|}$. Consider the vector $v' = \frac{v}{\|v\|}$. Then $\|v'\| = 1$, and some quick algebra reveals $x = \|Av'\|$. Therefore $x \in \omega$, so $\alpha \subseteq \omega$. This means the operator norm can be calculated by looking at either sup: they're the same set. □

Proposition 24. *The operator norm in Equation 34 is a norm on $\mathcal{L}(V, W)$. That is, it satisfies the norm axioms.*

Proof. We go through the norm axioms one by one.

(i) $\|A\| \geq 0$

The norm in W is always a non-negative quantity.

(ii) $\|A\| = 0$ iff $A = 0$

Let $v \in V$ with $\|v\| = 1$. Then $0v = 0$, and $\sup \frac{0}{1} = 0$ so $\|0\| = 0$. Now suppose $\|A\| = 0$. Then

$$\sup_{\|v\|=1} \|Av\| = 0$$

so that $\|Av\| \leq 0$ for each v. But norms of vectors in W are always ≥ 0, so $\|Av\| = 0$ for each vector. But a vector in W with 0 norm forces the vector to be the 0 vector, again by the truth of the norm axioms assumed on W. Therefore $Av = 0$ for each $v \in V$, so A is by definition the 0 operator.

(iii) $\|kA\| = |k|\|A\|$

Left to the reader. Follows from definition of sups and linearity.

(iv) $\|A + B\| \leq \|A\| + \|B\|$

Let $A, B \in \mathcal{L}(V, W)$. Then $A + B \in \mathcal{L}(V, W)$ because it's a subspace. Now, for any $v \in V$ with $\|v\| = 1$,

$$\|A\| \geq \|Av\| = \|[(A + B) - B]v\| \geq \|(A + B)v\| - \|Bv\|$$

so that $\forall v \in V$ with $\|v\| = 1$,

$$\|Bv\| \geq \|(A + B)v\| - \|A\|$$

Therefore the same is true in the sup:

$$\|B\| \geq \|A + B\| - \|A\|$$

from which the triangle inequality follows.

□

So we have a normed vector space $\mathcal{L}(V, W)$. Being normed does not guarantee completeness. Recall that completeness requires Cauchy sequences A_n to converge to some A in the space with respect to the operator norm. If we inspect Equation 34, how can we guarantee convergence of a Cauchy A_n? Well, when an operator acts on a vector it lands in W, so it turns out to suffice for just the codomain to be complete. The proof uses techniques which are reminiscent of the ℓ^2 completeness proof:

Theorem 2. *$\mathcal{L}(V, W)$ is a Banach space, for normed vector space V and Banach W.*

Proof. It has already been shown that $\mathcal{L}(V, W)$ is a vector space for normed V and normed W, so it is certainly true for normed V and Banach W. Therefore it remains to show completeness of $\mathcal{L}(V, W)$.

Let $A_n \in \mathcal{L}(V, W)$ be Cauchy. That is, $\forall \epsilon > 0, \exists N$:

$$\|A_n - A_m\| < \epsilon \,\text{whenever}\, n, m > N$$

We want to find an operator A that is the limit of this cauchy sequence, so we select a natural canditate. Define the operator $A : V \longrightarrow W$ by

$$Av = \lim A_n v$$

Is this even well defined? That is, does the limit of the $A_n v$'s always exist for each v? Let us check. Take any $v \in V$, nonzero. Then we know

$$\|(A_n - A_m)v\| \leq \|A_n - A_m\|\|v\|$$

But A_n is Cauchy, so $\forall \epsilon > 0, \exists N$:

$$\|A_n - A_m\| < \frac{\epsilon}{\|v\|} \,\text{whenever}\, n, m > N$$

Following the chain backwards we see therefore that the $A_n v$ sequence is Cauchy in W. But W is complete by assumption, so $\lim A_n v$ exists in W. Therefore A is well defined for all v. It remains to show that A is linear and bounded. Let us first show linearity. Let $v, w \in V$. Then

$$A(v) + A(w) = \lim A_n v + \lim A_n w = \lim A_n (v + w) = A(v + w)$$

Because each A_n and A_m is linear and the sum of the limits is the limit of the sum when each limit exists. Now let $\lambda \in \mathbb{C}$. Then :

$$A\lambda v = \lim A_n \lambda v = \lim \lambda A_n v = \lambda \lim A_n v = \lambda A v$$

because each A_n is linear and complex numbers can be pulled out of convergent sequences. It remains to show A is bounded. Observe that $\exists N$:

$$\|A_n - A_m\| < 1 \text{ whenever } n, m > N$$

which implies that for any v such that $\|v\| = 1$,

$$\|(A_n - A_m)v\| < 1 \text{ whenever } n, m > N$$

Now we may take the limit on m:

$$\|(A_n - A)v\| \leq 1 \text{ whenever } n > N$$

But this is true $\forall v$ on the unit sphere. This means $A_{N+1} - A$ is a bounded linear operator. But $\mathcal{L}(V, W)$ is a vector space, so we can multiply by -1 and add other vectors. So $A - A_{N+1} \in \mathcal{L}(V, W)$ and $A - A_{N+1} + A_{N+1} = A \in \mathcal{L}(V, W)$. Although we've coyly named this operator A and the sequence is A_n, we have not yet shown it is the limit. Let $\epsilon > 0$. Since A_n is Cauchy, $\exists N$:

$$\|A_n - A_m\| < \frac{\epsilon}{4} \text{ whenever } n, m > N$$

But the operator norm is a sup, so pick any $v \in V$ s.t. $\|v\| = 1$, and we have that

$$\|(A_n - A_m)v\| < \frac{\epsilon}{4} \text{ whenever } n, m > N$$

Taking the limit on m, we have:

$$\|(A_n - A)v\| \leq \frac{\epsilon}{4} < \frac{\epsilon}{2} \text{ whenever } n > N$$

But this is true $\forall v$ on the unit sphere. So taking the sup, we have

$$\|A_n - A\| \leq \frac{\epsilon}{2} < \epsilon \text{ whenever } n > N$$

Which is the statement $\lim A_n = A$. So $\mathcal{L}(V, W)$ is a complete. □

Corollary 3. *$\mathcal{L}(H)$ is a Banach space for any Hilbert space H.*

Proposition 25. *Let $A \in \mathcal{L}(V, W)$. Then for any convergent sequence $v_n \in V$, $A \lim v_n = \lim Av_n$.*

Proof. Define $v = \lim v_n$. Let $\epsilon > 0$. Then $\exists N$:

$$\|v_n - v\| < \frac{\epsilon}{\|A\|} \text{ whenever } n > N$$

So we now look at the image of the vector difference

$$\|Av_n - Av\| = \|A(v_n - v)\| \leq \|A\|\|v_n - v\| < \epsilon \text{ whenever } n > N$$

And so $Av = \lim Av_n$. □

Often, operators are not defined on the entire Hilbert space H, but will be defined on a dense subspace of H.

Definition 35. *An operator $A : D_A \longrightarrow H$ is said to be densely defined if D_A is a subspace of H and $\overline{D_A} = H$. That is, every element of H can be written as the limit of a sequence in D_A.*

Proposition 26. *Let $A : D_A \longrightarrow H$ be a densely defined bounded linear operator. Then for any $v_n \in D_A$ with $\lim v_n = v$ exists in H, $\lim Av_n$ exists, even if $v \notin D_A$.*

Proof. $\|v_n - v_m\|$ is Cauchy, so $\exists N$ such that it can be made less than $\frac{\epsilon}{\|A\|}$. Consider the sequence $Av_n \in H$. Therefore

$$\|Av_n - Av_m\| = \|A(v_n - v_m)\| \leq \|A\|\|v_n - v_m\| < \epsilon \,\text{whenever}\, n, m > N$$

□

Therefore Av_n is a Cauchy sequence in a complete space H, so that $\lim Av_n$ exists.

In particular, the position operator P and momentum operator Q will be defined only on dense subsets of the Hilbert space of wave functions, which we will discuss in detail in a later section. The question then arises: can either of these operators be extended by writing something like $Q\psi = \lim Q\psi_n$? The answer is no, because as we will show Q and P are not bounded and Proposition 26 can't be applied. The result is true for bounded operators, so we prove it now:

Theorem 3. *Let $A : D_A \longrightarrow H$ be a densely defined bounded linear operator. Then A can be uniquely extended to a bounded linear operator $\overline{A}$ defined on the whole space. That is, $\exists!\, \overline{A} \in \mathcal{L}(H) : A \subseteq \overline{A}$. Furthermore, $\|\overline{A}\| = \|A\|$.*

Proof. Since any $v \in H$ can be written as $v = \lim v_n$ for $v_n \in D_A$, we define $\overline{A} : H \longrightarrow H$ by $\overline{A}v = \lim Av_n$. By Proposition 26, $\lim Av_n$ exists. It remains to show $\overline{A}$ is linear and bounded. Let $v, w \in H$. Then

$$\begin{aligned}\overline{A}(v) + \overline{A}(w) &= \overline{A}\lim v_n + \overline{A}\lim w_n = \lim Av_n + \lim Aw_n = \lim Av_n + Aw_n \\ &= \lim A(v_n + w_n) = \overline{A}\lim(v_n + w_n) = \overline{A}(\lim v_n + \lim w_n) = \overline{A}(v + w) \quad (4.5)\end{aligned}$$

For scalar multiplication,

$$\begin{aligned}\overline{A}(kv) = \overline{A}(k\lim v_n) = \overline{A}(\lim kv_n) = \lim Akv_n = \lim kAv_n = k\lim Av_n \\ = k\overline{A}\lim v_n = k\overline{A}(v) \quad (4.6)\end{aligned}$$

where in each case the sequences v_n and w_n are in D_A and converge to v and w respectively. To see boundedness, let $v \in H, \|v\| = 1$. Then $v = \lim v_n$, and

$$\begin{aligned}\|\overline{A}v\| = \|\lim Av_n\| = \lim\|Av_n\| \leq \lim\|A\|\|v_n\| \\ = \|A\|\lim\|v_n\| = \|A\|\|\lim v_n\| = \|A\| \quad (4.7)\end{aligned}$$

Therefore for any v s.t. $\|v\| = 1$, $\|\overline{A}v\| \leq \|A\|$. Therefore the same is true in the sup: $\sup_{v\neq 0} \|\overline{A}v\| \leq \|A\|$, so $\overline{A}$ is bounded. However, the normalized $v \in D_A$ are also in H, so $\|A\| \leq \|\overline{A}\|$. Therefore $\|A\| = \|\overline{A}\|$. So we have an extension $\overline{A} \in \mathcal{L}(H)$ that has the same norm as the original operator. We wish to establish uniqueness. Suppose there were another extension $B \in \mathcal{L}(H)$. Then let $v \in H$. We can find $v_n \in D_A$ s.t. $v = \lim v_n$. Then by Proposition 25, $Bv = \lim Bv_n$. But B extends A, so $Bv_n = Av_n$ for each n. Therefore

$$Bv = \lim Bv_n = \lim Av_n = \overline{A}v.$$

Therefore $B = \overline{A}$. □

We conclude this section with a definition:

Definition 36. *A densely defined linear operator $A : D_A \longrightarrow H$ is said to be closed if the combination of statements*

$$\alpha_n \in D_A,\ \lim \alpha_n = \alpha \text{ exists},\ \lim A\alpha_n = \beta \text{ exists}$$

together imply that:

$$\alpha \in D_A,\ \text{and}\ A\alpha = \beta$$

Proposition 27. *Let $A : D_A \longrightarrow H$ be a closed operator. Then the operator $A - \lambda : D_A \longrightarrow H$ is closed and if A is 1-1, $A^{-1} : \text{ran}\, A \longrightarrow D_A$ is also closed.*

Proof. Let us first check closedness for $A-\lambda$. Take an $\alpha_n \in D_A$ with $\lim \alpha_n = \alpha$ exists and $\lim(A - \lambda)\alpha_n$ exists. Since A is closed $\alpha \in D_A$, and $\lim A\alpha_m = A\alpha$ so

$$\lim(A - \lambda)\alpha_n = \lim A\alpha_n - \lambda\alpha_n = A\alpha - \lambda\alpha = (A - \lambda)\alpha$$

since complex numbers can be pulled out of convergent sequences. Now let us check for A^{-1} for 1-1, closed A. Let $\omega_n \in \text{ran}\, A$ with $\lim \omega_n = \omega$ exists and $\lim A^{-1}\omega_n$ exists. Well for each n, $\exists \alpha_n \in D_A$ s.t. $\omega_n = A\alpha_n$. Then $\lim A^{-1}\omega_n = \lim \alpha_n =: \alpha$ exists, and $\lim \omega_n = \lim A\alpha_n$ exists. Since A is closed, $\alpha \in D_A$, and $A\alpha = \omega$. Therefore $\omega \in \text{ran}\, A$, and $A^{-1}\omega = \alpha$. □

4.1 Adjoints

The adjoint of an operator is akin to the complex conjugate of a number. In the finite dimensional case, the adjoint A^* of a matrix A can be defined as its transpose conjugate. In the case of a matrix $A : V \longrightarrow V$, some linear algebra verifies the fact that $\forall u, v \in V$:

$$\langle u, Av\rangle = \langle A^*u, v\rangle \tag{4.8}$$

In infinite dimensions, there is not an analagous way to immediately construct the adjoint of an operator. Instead, we start our definition with 4.8.

Definition 37. *Let $A : D_A \longrightarrow H$ be a densely defined linear operator. The adjoint A^* of A is defined by*

$$D_{A^*} = \left\{ \psi \in H \middle| \exists \eta \in H : \forall \alpha \in D_A, \langle \psi, A\alpha \rangle = \langle \eta, \alpha \rangle \right\} \tag{4.9}$$

And we define $A^*\psi = \eta$ for $\psi \in D_{A^*}$.

We wish to show well definition. Let $A : D_A \longrightarrow H$ be a densely defined operator and let $\psi \in D_{A^*}$. Suppose there is an η and a ζ that satisfy the conditions in 4.9, i.e. $A^*\psi = \eta$ and $A^*\psi = \zeta$. Then $\forall \alpha \in D_A$,

$$\langle \eta, \alpha \rangle = \langle \zeta, \alpha \rangle \implies \langle \eta - \zeta, \alpha \rangle = 0 \quad \forall \alpha \in D_A$$

Which means $\eta - \zeta$ is perpendicular to a dense subset of H. Therefore $\eta - \zeta = 0 \implies \eta = \zeta$ and well definition is established.

Definition 38. *An operator is said to be self-adjoint if $A = A^*$. That is, $D_A = D_{A^*}$ and $A\psi = A^*\psi$, $\forall \psi \in D_A$.*

A slightly weaker condition is for an operator to be symmetric:

Definition 39. *A densely defined operator $A : D_A \longrightarrow H$ is said to be symmetric if $\forall \alpha, \beta \in D_A$,*

$$\langle \alpha, A\beta \rangle = \langle A\alpha, \beta \rangle$$

The reason symmetry is not taken as the definition of self-adjointness is that D_A is not, in general, the whole Hilbert space H. It's usually just a dense subset. The adjoint of an operator may be defined on vectors outside its original domain.

Proposition 28. *A densely defined operator $A : D_A \longrightarrow H$ is symmetric iff $A \subseteq A^*$.*

Proof. Suppose A is symmetric. Then every $\alpha \in D_A$ satisfies the condition for being in D_{A^*}, and in each case $A\alpha = A^*\alpha$. Therefore $A \subseteq A^*$. Now suppose $A \subseteq A^*$. Then every $\alpha \in D_A$ satisfies the condition to be in D_{A^*}, and since $A^*\alpha = A\alpha$ for each $\alpha \in D_A$. Therefore A is symmetric. □

Corollary 4. *Self adjoint operators are symmetric.*

Definition 40. *An operator is said to be essentially self-adjoint if A^{**} is self-adjoint.*

4.2 Ranges and Kernels

Definition 41. *The range of an operator $A : D_A \longrightarrow H$ is defined by*

$$ran\, A := \left\{ \beta \in H \middle| \beta = A\alpha \, for \, some \, \alpha \in D_A \right\}$$

That is, the range of an operator is the set of vectors that the domain gets mapped to.

Remark 8. *ran* $A \subseteq H$. *That is, the range is a subset of the codomain.*

Definition 42. *The null space (or kernel) of an operator* A *is the defined by:*

$$N(A) = ker\, A := \left\{ \alpha \in D_A \middle| A\alpha = 0 \right\}$$

Remark 9. $N(A) \subseteq D_A$. *That is, the null space is a subset of the domain. It is the preimage of the* 0 *vector.*

Proposition 29. $(ran\, A)^{\perp} = N(A^*)$

Proof. Let $\beta \in (\text{ran}\, A)^{\perp}$. Then $\langle \beta, A\alpha \rangle = 0$ and $\beta \in D_{A^*}$ with $A^*\beta = 0$, so $\beta \in N(A^*)$. The converse is the same thing backwards and is left to the reader. □

Proposition 30. $(A - \lambda)^* = A^* - \overline{\lambda}$

Proof. Left to the reader. □

Remark 10. $N(A - \lambda) = E_\lambda$ *for* $\lambda \in \Lambda(A)$.

4.3 Projectors

Projectors show up in quantum mechanics to describe measurement results. In essence, when an eigenvalue λ is measured for a system in the state ψ, the state after measurement is the piece of ψ that was in the eigenspace corresponding to λ. The reality is a little more complicated, as states are not vectors, but this is the spirit of the process. The important intuition is that measuring λ changes the state by projecting it onto the eigenspace.

(picture).

Definition 43. *A linear operator* $P \in \mathcal{L}(H, H)$ *is said to be a projector if:*

(i) $P^2 = P$

(ii) $P* = P$

Proposition 31. *Let* $P : \mathcal{H} \longrightarrow \mathcal{H}$ *be a projector. For any* $a \in ran\,(P)$, $Pa = a$.

Proof. $a = P\psi$ for some $\psi \in \mathcal{H}$. Now, hit the equation with P on both sides.

$$Pa = P^2\psi = P\psi = a$$

□

Proposition 32. *Let* $P : \mathcal{H} \longrightarrow \mathcal{H}$ *be a projector. Then* $\|P\| \leq 1$.

Proof. Denote $M = \text{ran}\,(P)$. Let $\psi \in H$. Note that $\psi = P\psi + (\psi - P\psi)$. We will denote $\phi = P\psi$ and $\eta = \psi - P\psi$. Observe that $P\eta = P\psi - P\psi = 0$. Now,

$$\|P\psi\| = \|P(\phi + \eta)\| = \|P\phi + P\eta|Vert = \|\phi\| \leq \|\phi + \eta\| = \|\psi\|$$

so $\|P\| \leq 1$. □

There is something very important about projectors:

Proposition 33. *Given a projector $P : \mathcal{H} \longrightarrow \mathcal{H}$, the space $M = ran(P)$ is a closed subspace.*

Proof. Let a be a limit point of M. Then $a = \lim a_n = \lim P\psi_n$ for $\psi_n \in H$. But $P\psi_n = Pa_n$, so $a = \lim Pa_n$. Now, since P is bounded and its domain is all of H, $\lim Pa_n = P \lim a_n = Pa$. so $a = Pa \in \text{ran}\,(P)$. □

4.4 Unitary Maps

Definition 44. *A linear map $U \in \mathcal{L}(H, H)$ is said to be unitary if:*

(i) $ran(U) = H$

(ii) $\forall f, g \in H, \langle Uf, Ug\rangle = \langle f, g\rangle$

Proposition 34. *For unitary maps, $U^* = U^{-1}$.*

4.5 Spectra

The spectrum of an observable O tells you what possible values you can get when you measure that observable. For example, consider the measurement of a particle along the x-axis. The spectrum of the x-position operator will be shown to be the whole real line, so that the particle can be found anywhere. It is frequently (incorrectly) stated that these are eigenvalues, and that the delta functions are eigenvectors. This notion can be made more precise, but we do not explore that idea here. It will be shown that the position operator has no eigenvalues, and therefore has no eigenfunctions. Delta functions are in fact not even functions. We require a more general notion of the spectrum of an operator, which contains more than just eigenvalues.

Definition 45. *A number $\lambda \in \mathbb{C}$ is called an eigenvalue of a linear operator A if $\exists v \in D_A, v \neq 0$ s.t.*

$$Av = \lambda v \tag{4.10}$$

In this case, v is called the eigenvector corresponding to λ.

Definition 46. *We write $\Lambda(A)$ to denote the set of eigenvalues of A.*

Definition 47. *We write E_λ to denote the set of eigenvectors corresponding to λ.*

Proposition 35. *E_λ is a subspace of D_A for each eigenvalue λ.*

Proof. Left to the reader. □

Linear operators act like multiplication, but in this case, eigenvectors really see operators as multiplication by a number.

Proposition 36. *The eigenvalues of a symmetric operator are real.*

Proof. Let $A : D_A \longrightarrow H$ be a symmetric operator. Let $\lambda \in \Lambda(A)$. Then $\exists v \neq 0$: $Av = \lambda v$. We examine the quantity

$$\langle v, Av\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle$$

Since A is symmetric, we have

$$\langle v, Av\rangle = \langle Av, v\rangle$$

and similar to above

$$\langle Av, v\rangle = \langle \lambda v, v\rangle = \overline{\lambda}\langle v, v\rangle$$

And therefore

$$\overline{\lambda}\langle v, v\rangle = \lambda\langle v, v\rangle$$

But $v \neq 0$, so $\lambda = \overline{\lambda}$. Therefore $\Lambda(A) \subseteq \mathbb{R}$. □

Corollary 5. *The eigenvalues of a self-adjoint operator are real.*

Proposition 37. *Given densely defined $A : D_A \longrightarrow H$, the operator $A - \lambda : D_A \longrightarrow H$ is $1-1$ iff $\lambda \notin \Lambda(A)$.*

Proof. Suppose $\lambda \in \Lambda(A)$. Then $\exists v \neq 0$: $Av = \lambda v$. Therefore $(A-\lambda)v = 0$ and $v \neq 0$, so $A - \lambda$ is not $1-1$. Therefore if $A - \lambda$ is $1-1$, λ can't be in $\Lambda(A)$. Now suppose $\lambda \notin \Lambda(A)$. Then the equation $Av = \lambda v$ forces $v = 0$. Therefore $(A-\lambda)v = 0 \implies v = 0$. Therefore A is $1-1$. □

It is clear that when $\lambda \in \Lambda(A)$, the operator A looks exactly like λ to the vectors in E_λ. So the operator $A - \lambda$ gets all messed up: to some vectors it looks like the 0 operator, and therefore $(A-\lambda)^{-1}$ does not exist. So we give a name to the set of λ where $A - \lambda$ is a very nice operator:

Definition 48. *The resolvent set of an operator A is defined by:*

$$\rho(A) := \left\{\lambda \in \mathbb{C} : (A-\lambda)^{-1} \text{ exists and is in } \mathcal{L}(H)\right\} \tag{4.11}$$

Definition 49. *The spectrum of an operator A is defined by:*

$$\sigma(A) := \rho(A)^c \tag{4.12}$$

Any densely defined operator A thus splits up the complex plane into numbers that look A (the spectrum) and numbers that don't (the resolvent set).

Corollary 6. $\Lambda(A) \subseteq \sigma(A)$

But how can an operator A look like a number λ without the λ being an eigenvalue? Well, according to our definition, if it's not an eigenvalue but is in the spectrum, then $(A-\lambda)^{-1}$ exists, but either its domain is not all of H or it isn't bounded. As we will show, both can't happen at the same time.

Proposition 38. *For self-adjoint $A : D_A \longrightarrow H$, $\lambda \in \Lambda(A)$ iff $\overline{ran\,(A-\lambda)} \neq H$*

Proof. Suppose λ is an eigenvalue of A. Then

$$\overline{\operatorname{ran}(A-\lambda)} = \operatorname{ran}(A-\lambda)^{\perp\perp} = N([A-\lambda]^*)^{\perp} = N(A^*-\overline{\lambda})^{\perp} = N(A-\lambda)^{\perp} \quad (4.13)$$

But λ is an eigenvalue, so $\exists v \neq 0 \in N(A-\lambda)$, so $N(A-\lambda) \supset 0$. Therefore

$$\overline{\operatorname{ran} A-\lambda} = N(A-\lambda)^{\perp} \neq H$$

The prove of the converse is essentially the same steps backwards and is left to the reader. □

Proposition 39. *Let $A : D_A \longrightarrow H$ be a densely defined operator. Then if $\lambda \notin \Lambda(A)$ and $\lambda \in \sigma(A)$, then either $(A-\lambda)^{-1}$ is unbounded or $ran\,A \neq H$, but not both.*

Proof. Suppose $\lambda \in \sigma(A)$ but is not an eigenvalue. Then one of the two conditions stated have to be true, by the definition of the spectrum. Suppose both were true. Then we would have a bounded linear operator defined on a dense subset of H, and we can extend it to H. Then λ would be in the resolvent set, a contradiction. □

We characterize the spectrum of an operator into two types:

Definition 50. *The point spectrum $\sigma_p(A)$ is defined as:*

$$\sigma_p(A) := \left\{ \lambda \in \mathbb{C} \middle| \overline{ran\,(A-\lambda)} \neq H \right\}$$

Definition 51. *The continuous spectrum $\sigma_c(A)$ is defined as:*

$$\sigma_c(A) := \left\{ \lambda \in \mathbb{C} \middle| ran\,(A-\lambda) \neq \overline{ran\,(A-\lambda)} \right\}$$

The following theorem is quintessential to quantum mechanics: it's why your arms aren't $2+3i$ feet long. The proof is based on one from [6], but the second half is my own and avoids their use of proof by contradiction.

Theorem 4. *Self-adjoint operators have real spectra.*

Proof. Let $A : D_A \longrightarrow H$ be a densely defined self-adjoint operator. Let $\lambda \in \mathbb{C}$ be a complex number that is not real, i.e. $\lambda = \xi + i\eta$, with $\eta \neq 0$. Then $\lambda \notin \Lambda(A)$. Therefore $A-\lambda$ is 1-1, so $(A-\lambda)^{-1} : \operatorname{ran} A \longrightarrow H$ exists. Let $g \in \operatorname{ran}(A-\lambda)$. Then $\exists f \in D_A$: $g = (A-\lambda)f$. We examine the quantity

$$\begin{aligned} \|g\|^2 &= \langle (A-\lambda)f, (A-\lambda)f\rangle = \langle (A-\xi)f - i\eta f, (A-\xi)f - i\eta f\rangle \\ &= \langle (A-\xi)f, (A-\xi)f\rangle + \langle (A-\xi)f, -i\eta f\rangle + \langle -i\eta f, (A-\xi)f\rangle + \langle -i\eta f, -i\eta f\rangle \\ &= \|(A-\xi)f\|^2 + \|\eta f\|^2 - i\eta\langle (A-\xi)f, f\rangle + i\eta\langle f, (A-\xi)f\rangle \end{aligned} \quad (4.14)$$

But A is self-adjoint, and ξ is a real number, so $A - \xi$ is self-adjoint. Therefore the cross terms cancel and we have

$$\|g\|^2 = \|(A-\xi)f\|^2 + \|\eta f\|^2 \geq |\eta|^2\|f\|^2$$

So that $\|g\| \leq |\eta|\|f\|$. But since f is the vector g came from, $\|g\| \leq |\eta|\|(A - \lambda)^{-1}g\|$, so that $\forall g \in \operatorname{ran} A$,

$$\frac{\|(A-\lambda)^{-1}g\|}{\|g\|} \leq \frac{1}{|\eta|}$$

Since $\eta \neq 0$, and this is true $\forall g$, $\|(A-\lambda)^{-1}\|$ is bounded above by $\frac{1}{|\eta|}$ and $(A - \lambda)^{-1}$ is a bounded linear operator. Since λ is not an eigenvalue, $\overline{\operatorname{ran}(A-\lambda)} = H$, and $A - \lambda$ is 1-1. Since A is self-adjoint, it is closed, so we have a 1-1 closed operator and therefore $(A-\lambda)^{-1}$ is closed. Now let $g \in H$. Then $\exists g_n \in \operatorname{ran} A - \lambda$ with $\lim g_n = g$. By 26, $\lim(A-\lambda)^{-1}g_n$ exists. Therefore we have all the initial conditions for a closed operator, and one of the results is that $g \in \operatorname{ran}(A-\lambda)$. Therefore $H \subseteq \operatorname{ran}(A = \lambda)$. But $\operatorname{ran}(A-\lambda) \subseteq H$ by definition of A. So $\operatorname{ran}(A-\lambda)^{-1} = H$. Therefore $(A-\lambda)^{-1}$ exists and is in $\mathcal{L}(H)$.

That means that if λ has any imaginary part, it has to be in the resolvent set. Therefore $\sigma(A) \subseteq \mathbb{R}$. □

Proposition 40. *The eigenvalues of a unitary operator U lie on the unit circle in $\mathbb{C}$. That is, for any $\lambda \in \Lambda(U)$, $|\lambda| = 1$.*

Proof. Let $U : H \longrightarrow H$ be a unitary operator. Let $\lambda \in \Lambda(U)$. Then $\exists f \neq 0$: $Uf = \lambda f$. Without loss of generality assume $\|f\| = 1$.

$$1 = \langle f, f\rangle = \langle Uf, Uf\rangle\langle \lambda f, \lambda f\rangle = |\lambda|^2\langle f, f\rangle = |\lambda|^2$$

So λ is on the unit circle. □

It is also true that $\sigma(U)$ lies on the unit circle, but we do not prove that here.

Proposition 41. *If an operator $B : D_B \longrightarrow H$ is self-adjoint and its inverse $B^{-1} : H \longrightarrow D_B$ exists in $\mathcal{L}(H)$, then B^{-1} is self-adjoint.*

Proof.

$$D_{B^{-1*}} = \left\{\psi \in H \middle| \exists \eta \in H : \forall \alpha \in D_{B^{-1}}, \langle \psi, B^{-1}\alpha\rangle = \langle \eta, \alpha\rangle\right\}$$

Take an $\psi \in H$. $H = \operatorname{ran} A$, so $\psi = B\phi$ for some $\phi \in D_B$, and therefore $B^{-1}\psi = \phi$. Therefore

$$\langle \psi, B^{-1}\alpha\rangle = \langle B\phi, B^{-1}\alpha\rangle = \langle \phi, BB^{-1}\alpha\rangle = \langle \phi, \alpha\rangle$$

So $D_{B^{-1*}} = H$ and $B^{-1*}\phi = B^{-1}\phi$. This completes the proof.

□

4.6 The Trace

Proposition 42. *For any ON basis e_n, and for any $\phi, \psi \in H$, we have*

$$\langle \phi, \psi \rangle = \sum_{n=1}^{\infty} \langle \phi, e_n \rangle \langle e_n, \psi \rangle$$

Proof. Since e_n is an ON basis, we have

$$\phi = \sum_{n=1}^{\infty} \langle e_n, \phi \rangle e_n$$

and

$$\psi = \sum_{m=1}^{\infty} \langle e_m, \psi \rangle e_m$$

so that

$$\begin{aligned} \langle \phi, \psi \rangle &= \langle \sum_{n=1}^{\infty} \langle e_n, \phi \rangle e_n, \sum_{m=1}^{\infty} \langle e_m, \psi \rangle e_m \rangle \\ &= \langle \lim_{N\to\infty} \sum_{n=1}^{N} \langle e_n, \phi \rangle e_n, \lim_{M\to\infty} \sum_{m=1}^{M} \langle e_m, \psi \rangle e_m \rangle \\ &= \lim_{N\to\infty} \lim_{M\to\infty} \langle \sum_{n=1}^{N} \langle e_n, \phi \rangle e_n, \sum_{m=1}^{M} \langle e_m, \psi \rangle e_m \rangle \end{aligned} \quad (4.15)$$

But now we only have finite sums in the inner product, so they can be pulled out:

$$\begin{aligned} &= \lim_{N\to\infty} \lim_{M\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} \langle \langle e_n, \phi \rangle e_n, \langle e_m, \phi \rangle e_m \rangle \\ &= \lim_{N\to\infty} \lim_{M\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} \langle \phi, e_n \rangle \langle e_m, \psi \rangle \langle e_n, e_m \rangle \\ &= \lim_{N\to\infty} \sum_{n=1}^{n} \sum_{m=1}^{M} \langle \phi, e_n \rangle \langle e_n, \psi \rangle = \sum_{n=1}^{\infty} \langle \phi, e_n \rangle \langle e_n, \psi \rangle \end{aligned} \quad (4.16)$$

□

Remark 11. *The shorthand for this is to just write put $I = \sum_{n=1}^{\infty} |e_n\rangle\langle e_n|$ between $\langle \phi|$ and $|\psi\rangle$ in the inner product $\langle \phi|\psi\rangle$.*

The trace of an operator is an important tool. Before defining the trace, we make a preliminary definition.

Definition 52. *We define a bounded linear operator $A : H \longrightarrow H$ to be trace-class if the quantity*

$$\sum_{n=1}^{\infty} \langle e_n, Ae_n \rangle$$

is finite for any ON basis e_n. For a trace-class operator A, we define the trace of A to be

$$Tr(A) := \sum_{n=1}^{\infty} \langle e_n, Ae_n \rangle$$

for an ON basis labeled $\{e_n\}$.

Proposition 43. *The trace is basis independent; i.e. the trace is well defined.*

Proof. Let e_n and f_m be two ON bases. We wish to show

$$\sum_{n=1}^{\infty} \langle e_n, Ae_n \rangle = \sum_{n=1}^{\infty} \langle f_m, Af_m \rangle$$

We will denote the LHS by $Tr_e(A)$ an the RHS by $Tr_f(A)$.

Since e_n is an ON basis, we know that each f_m can be written as

$$f_m = \sum_{n=1}^{\infty} \langle e_n, f_m \rangle e_n$$

So that

$$\begin{aligned} Tr_f(A) &= \sum_{m=1}^{\infty} \Big\langle \sum_{i=1}^{\infty} \langle e_i, f_m \rangle e_i, A \sum_{j=1}^{\infty} \langle e_j, f_m \rangle e_j \Big\rangle \\ &= \sum_{m=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle e_i, f_m \rangle \langle f_m, e_i \rangle \langle e_i, Ae_j \rangle \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sum_{m=1}^{\infty} \langle e_j, f_m \rangle \langle f_m, e_i \rangle \langle e_i, Ae_j \rangle \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle e_i, Ae_j \rangle \sum_{m=1}^{\infty} \langle e_j, f_m \rangle \langle f_m, e_i \rangle \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle e_i, Ae_j \rangle \delta_{ij} = \sum_{i=1}^{\infty} \langle e_i, Ae_i \rangle = Tr_e(A) \end{aligned} \tag{4.17}$$

□

Definition 53. *Given a bounded linear map $A : H \longrightarrow H$, A is said to be positive if $\forall f \in H$,*

$$\langle f, Af \rangle \geq 0$$

Definition 54. *States of a system are said to be pure if there exists $\psi \in H$ such that*

$$\rho\phi = \langle \psi, \phi \rangle \psi$$

In this case, we denote $\rho = \rho_\psi$.

Proposition 44. *Given a pure state ρ_ψ and self-adjoint operator A,*

$$Tr(\rho_\psi A) = \langle \psi, A\psi \rangle$$

Proof.

$$Tr(\rho_\psi A) = \sum \langle e_i, \rho_\psi A e_i \rangle = \sum \langle e_i, \langle \psi, A e_i \rangle \psi \rangle = \sum \langle A\psi, e_i \rangle \langle e_i, \psi \rangle = \langle A\psi, \psi \rangle = \langle \psi, A\psi \rangle \quad (4.18)$$

□

4.7 The Commutator

We finish the section with a brief discussion of the commutators of operators. Many of the definitions here come from future sections.

Definition 55. *For $A, B \in \mathcal{L}(\mathcal{H})$, we define the commutator of A and B to be*

$$[A, B] := AB - BA$$

Definition 56. *We say two bounded operators $A, B \in \mathcal{L}(\mathcal{H})$ commute if*

$$[A, B] = 0$$

Definition 57. *If A and B are unbounded self-adjoint operators, then we say A and B commute if*

$$[e^{itA}, e^{isB}] = 0$$

for some $t, s \in \mathbb{R}$.

Proposition 45. *On the Schwartz space $\mathcal{S}(\mathbb{R})$, it holds true that $\forall \psi \in \mathcal{S}$,*

$$QP\psi - PQ\psi = i\hbar\psi$$

Proof. Can be found in [1]. □

5. MEASURE THEORY

We wish to talk about the Hilbert space of square-integrable functions. However, the way we define "integrable" requires definitions from measure theory. We will also need projection-valued measures, a generalization of the notion of a measure, to talk about quantum measurement. The first type of measure we wish to talk about is a way to assign a length to sets of real numbers, so that we can integrate over them.

Definition 58. *A σ-algebra σ for a set M is a collection of subsets of M satisfying three axioms:*

(i) $\varnothing \in \sigma$

(ii) $E \in \sigma \implies E^c \in \sigma$

(iii) $E_1, E_2, ... \in \sigma \implies \bigcup_{n=1}^{\infty} E_n \in \sigma$

The subsets $E \subseteq M$ are said to be measurable sets.

Proposition 46. *Let E be a collection of subsets of a set M. The set σ defined by*

$$\sigma = \bigcap_{E \in \Omega,\, \sigma - alg} \Omega$$

is also a sigma algebra. This is also called the smallest sigma algebra containing E, or the σ-algebra generated by E, and is written $\sigma(E)$.

Proof. Can be found in [8]. □

Proposition 47. *If σ is a $\sigma - algebra$ for M, then there exists an E such that $\sigma = \sigma(E)$.*

Proof. Can be found in [8]. □

Definition 59. *A measure is a function $\mu : \sigma \longrightarrow [0, \infty]$ satisfying two axioms:*

1. $\mu(\emptyset) = 0$
2. $A_1, A_2, ... \in \sigma$ *with* $A_i \bigcap A_j = \emptyset$ *for* $i \neq j$ *(pairwise disjoint) implies that*

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

5.1 Measuring $\mathbb{R}$

Our primary task is to define a measure-based integral of complex number valued wave-functions. Our first task, then, is to define a measure of line segments.

Take the line segment $(3, 7)$. Clearly, if I were to measure the length of this segment, I would find it to be $7 - 3 = 4$.

Definition 60. *The length of an open interval $I = (a, b)$ is defined to be*

$$\ell(I) = b - a$$

In addition, we define that $\ell([a, b)) = \ell([a, b]) = \ell((a, b]) = b - a$.

Now consider a more complicated set, like $A = (2, 3) \cup (4, 5)$. Clearly the size of A should be 2, but can we relate this to the length by a definition? What if A were an extremely wonky set, like the set $C =$ all rational multiple of e between 3 and 5.3? Well, if I a roll of tape with numbers marking length, I could place pieces of tape over small subsets of C. If I kept doing this until I used enough tape to cover all of C, then the length of C is certainly no more than the total length of tape I used. First we define a rigorous generalization of our tape strategy:

Definition 61. *Given a set E, an open cover $\mathcal{O}$ for E is a collection of open subsets $\mathcal{O} = \{\mathcal{O}_\alpha\}_{\alpha \in A}$ such that $\bigcup_{\alpha \in A} \mathcal{O}_\alpha \supseteq E$, where A is an indexing set for $\mathcal{O}$. That is, an open cover for E is a collection of open sets that cover all of E.*

In particular, one may form open covers which are sequences of intervals, like our tape example. Since we know the length of each piece of tape, then the size of C should be how much tape we need at the very least to cover all of C. This is called the Lebesgue outer measure.

Definition 62. *The Lebesgue outer measure $\mu^* : \mathcal{P}(\mathbb{R}) \longrightarrow [0, \infty]$ by the equation*

$$\mu^*(E) = \inf \Big\{ \sum_{k=1}^{\infty} \ell(I_k) \Big| \bigcup_{k=1}^{\infty} I_k \supseteq E \Big\}$$

And each sequence I_k is understood to be a sequence of intervals.

This is an extremely clever way to assign "lengths" to sets which are much more unwieldy than ordinary intervals. However, in order for our definition to be sensible we must verify that the outer measure of a set agrees with its length. First, we state a useful theorem without proof:

Theorem 5. *(Heine-Borel) Any open cover for a closed, bounded subset of the real axis contains a finite subcover. That is, given a closed, bounded $M \subseteq R$, then for any open cover $\{\mathcal{O}_\alpha\}_{\alpha \in A}$ of M, there exists a finite subcover $\mathcal{O}_1, \mathcal{O}_2, ..., \mathcal{O}_n$ such that each $\mathcal{O}_i$ is one of the open sets in the cover.*

Proposition 48. *$\mu^*(I) = \ell(I)$ for open intervals I.*

Proof. First, we show $\mu^*(I) \leq \ell(I)$. Pick the sequence $I, \emptyset, \emptyset, ...$ This is an open cover of I, so $\sum_{k=1}^{\infty} I_k = \ell(I) + \ell(\emptyset) + ... = \ell(I)$ is in the set that μ^* takes the inf over. Therefore $\mu^*(I) \leq \ell(I)$.

The converse is much trickier. To do this, we write $I = (a, b)$ concretely. Let $\epsilon > 0$. Let $I^\epsilon = [a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2}] \subseteq (a, b)$. Let I_k be any open cover for I. Then the sequence I_k is also an open cover of I^ϵ. Now, by Theorem 5, we know that there is a finite subcover $J_1, J_2, ..., J_n$ of I_k for I^ϵ such that each J_i was an I_k for some k. So

$$\sum_{k=1}^{\infty} \ell(I_k) \geq \sum_{j=1}^{n} \ell(J_j)$$

But if you arrange the J_n in the right way it is easy to see that

$$\sum_{j=1}^{n} \ell(J_j) > b - a - \epsilon$$

But for any $\epsilon > 0$, equality might hold in the infimum over all open interval covers:

$$\inf \sum_{k=1}^{\infty} \ell(I_k) \geq b - a - \epsilon$$

And the same must hold $\forall \epsilon > 0$, so that

$$\inf \sum_{k=1}^{\infty} \ell(I_k) \geq b - a$$

But the left hand side is $\mu^*(I)$, and the right hand side is $\ell(I)$. This completes the proof. □

Definition 63. *A set E is said to be measurable if, for all $A \subseteq R$,*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

Proposition 49. *If E is measurable, then E^c is measurable.*

Proof. Follows immediately from the definition and the fact that $E^{cc} = E$. □

Proposition 50. *$\emptyset$ and $\mathbb{R}$ are measurable.*

Proof. First we show that $\emptyset$ is measurable. Let $A \subseteq \mathbb{R}$. Then

$$\mu^*(A \cap \emptyset) + \mu^*(A \cap \emptyset^c) = \mu^*(\emptyset) + \mu^*(A \cap \mathbb{R}) = \mu^*(A)$$

Then by Proposition 49 $\emptyset^c = \mathbb{R}$ is measurable. □

Proposition 51. *If $E_1, E_2, ...$ is a countable sequence of measurable sets then $\cup_i E_i$ is measurable.*

Proof. Can be found in [8]. □

Corollary 7. *The collection of measurable sets of* $\mathbb{R}$ *is a* σ*-algebra.*

Proposition 52. *If* $E \subseteq F$*, then* $\mu^*(E) \leq \mu^*(F)$*.*

Proof. Let I_k be any sequence of open intervals that covers F. Then the open cover I_k also covers E, because $E \subseteq F$. Therefore the set of open covers of E is a subset of open covers of F, so that the inf over E is less than or equal to the inf over F. i.e., $\mu^*(E) \leq \mu^*(F)$. □

Proposition 53. *Given measurable* $E, F \subseteq \mathbb{R}$*, the following equality holds:*

$$\mu^*(E \cup F) + \mu^*(E \cap F) = \mu^*(E) + \mu^*(F)$$

Proof.

$$\mu^*(E \cup F) + \mu^*(E \cap F) = \mu^*[(E \cup F) \cap F] + \mu^*[(E \cup F) \cap F^c] + \mu^*(E \cap F) = \mu^*(F) + \mu^*(E \cap F^c) + \mu^*(E \cap F) = \mu^*(E) + \mu^*(F) \quad (5.1)$$

□

Corollary 8. $\mu^*(E \cup F) \leq \mu^*(E) + \mu^*(F)$

Proposition 54. *If* $E_1, E_2, \ldots$ *is a countable sequence of disjoint measurable sets then*

$$\mu^*\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{i=1}^{\infty} \mu(E_i)$$

Proof. Can be found in [8]. □

Corollary 9. *The restriction of the Lebesgue outer measure to the set of measurable sets is a measure, and we denote it by* $\mu : \sigma \longrightarrow [0, \infty]$*.*

Proposition 55. *Let* $\mu : \sigma \longrightarrow [0, \infty]$ *be a measure. Let* $F \in \sigma$*. Then the function* $\mu|_F : \sigma \longrightarrow [0, \infty]$ *defined by*

$$\mu|_F(E) = \mu(E \cap F)$$

is also a measure.

Proof. We verify both measure axioms.

(i) Claim: $\mu|_F(\emptyset) = 0$
Proof: $\mu|_F(\emptyset) = \mu(\emptyset \cap F) = \mu(\emptyset) = 0$.

(ii) Claim: $\mu|_F(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu|F(E_i)$
Proof:

$$\mu|_F(\bigcup_{i=1}^{\infty} E_i) = \mu(F \cap \bigcup_{i=1}^{\infty} E_i) = \mu(\bigcup_{i=1}^{\infty} F \cap E_i) = \sum_{i=1}^{\infty} \mu(F \cap E_i) = \sum_{i=1}^{\infty} \mu|_F(E_i)$$

□

5.2 Measure-based integration

Integration is, among physicists, an "uncountably" infinite sum of infinitesimal objects. Mathematically, however, this is done with limits. In the case of real valued functions defined on the real axis, Riemann integration is the most obvious approach. It allows for easy calculation with antiderivatives according to the fundamental theorem of calculus. In order to see the connection between Riemann integration and measure theory, and in particular how the Riemann integral generalizes to the Lebesgue integral, we first precisely define the Riemann integral. First, we need to break up the interval to make all of the little rectangles.

Definition 64. *Given $a, b \in \mathbb{R}$, we define a partition $\mathcal{P}$ of the interval $[a, b]$ to be any ordered $n+1$-tuple of real numbers $(x_0, x_1, ..., x_n)$ satisfying $a = x_0 < x_1 < ... < x_n = b$. We define the symbol $\Delta x_i := x_i - x_{i-1}$.*

Definition 65. *We define the norm of a partition $\mathcal{P}$ to be*

$$\|P\| = \max \Delta x_i$$

over all of the Δx_i's in that partition.

Definition 66. *Given a partition $\mathcal{P} = (x_0, x_1, ..., x_n)$, we define a tagged partition $\dot{\mathcal{P}}$ to be any set of ordered pairs $([x_{i-1}, x_i], t_i)$ such that $t_i \in [x_{i-1}, x_i]$*

Now, if we have any function $f : [a, b] \longrightarrow \mathbb{C}$, we can add up the "area" of the rectangles generated by a given partition and the function to approximate the "area" under the function.

Definition 67. *Given a tagged partition $\dot{\mathcal{P}}$, we define the Riemann sum of f with respect to $\dot{\mathcal{P}}$ to be:*

$$S(f, \dot{\mathcal{P}}) = \sum_{i=0}^{n} f(t_i)\Delta x_i$$

Definition 68. *Let $f : [a, b] \longrightarrow \mathcal{C}$. We say that f is integrable if $\exists L$ such that $\forall \epsilon > 0$, $\exists \delta > 0$:*

$$\|S(f, \dot{\mathcal{P}}) - L\| < \epsilon \, whenever \|P\| < \delta$$

In other words,

$$\lim_{\|\dot{\mathcal{P}}\| \to 0} S(f, \dot{\mathcal{P}}) = L$$

In this case, we define the Riemann integral of f over $[a, b]$ by

$$\int_a^b f(x)dx := L$$

The use of the partition makes it easy to see how differentiation is useful in integration of functions with continuous derivatives. If $f = \frac{dF}{dx}$ for some F, then if we add up a lot of $f(t)dx$'s we get the sum of all the dF's, obtaining that the

integral is $F(b) - F(a)$. However, take the function $\phi : [0,1] \longrightarrow \mathcal{R}$ defined by $\phi(x) = 1$ if $x \in \mathbb{Q}$ and $\phi(x) = 0$ otherwise. Consider also the function ψ defined on the same interval but is instead 0 on the rationals and 1 on the irrationals. What is the area under these functions? The Riemann integral has no way to interpret such functions. The Lebesgue integral, however, does. The reason is that instead of working with just partitions, the idea is that you're allowed to break up the set you're integrating over in any way that you want.

Definition 69. *We say a function $s : \mathbb{R} \longrightarrow \mathbb{R}$ is simple if it has a finite range, i.e. $s(\mathbb{R}) = \{s_1, ..., s_n\}$. We denote the preimage of s_i by A_i.*

Remark 12. *ψ and ϕ as defined above are simple.*

Definition 70. *The Lebesgue integral of a simple function s with respect to the measure μ is defined by:*

$$\int_{\mathbb{R}} s \, d\mu := \sum_{i=1}^{n} s_i \mu(A_i)$$

Compare this with Definition 67.

Proposition 56. *$\int \phi \, d\mu = 0$ and $\int \psi \, d\mu = 1$*

Proof. $\int \phi \, d\mu = 1\mu(\mathbb{Q}) + 0\mu(\mathbb{I}) = 0$. Similarly for ψ. □

In the case that we don't want to integrate over the whole real line, we can restrict the integration domain to just a measurable subset.

Definition 71. *Let s be a simple function and Ω a measurable subset. The integral of s over Ω is defined by:*

$$\int_{\Omega} s \, d\mu := \sum_{i=1}^{n} s_i \mu(A_i \cap \Omega)$$

Definition 72. *Let $f : \mathbb{R} \longrightarrow [0, \infty]$ be a nonnegative measurable function. The Lebesgue integral of f is defined by the equation:*

$$\int f \, d\mu := \sup_{0 \leq s \leq f} \int s \, d\mu$$

A nonnegative measurable f is said to be integrable if $\int f \, d\mu$ is finite.

Definition 73. *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a measurable function. A measurable f is said to be integrable if $|f|$ is integrable.*

Definition 74. *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be an integrable function. The Lebesgue integral of f is defined by the equation:*

$$\int f \, d\mu := \int f_+ \, d\mu + \int f_- \, d\mu$$

where f_+ and f_- are defined by

$$f_+(x) = max(f(x), 0) \qquad and \qquad f_-(x) = max(-f(x), 0)$$

That is, f_+ is positive when f is positive and 0 when f is negative, and f_- is negative when f is negative and 0 when f is positive. (include picture)

Definition 75. *Let $f : \mathbb{R} \longrightarrow \mathbb{C}$ be a measurable function. f is said to be integrable if $|f|$ is integrable.*

Definition 76. *Let $f : \mathbb{R} \longrightarrow \mathbb{C}$ be integrable. The Lebesgue integral of f is defined by the equation:*

$$\int f \, d\mu = \int Re(f) \, d\mu + \int Im(f) \, d\mu$$

Remark 13. *For complex valued functions, the Lebesgue integral is defined by the real-valued real and imaginary component functions. The integrals of the real valued functions are defined by the nonnegative f_+ and f_- functions, whose integrals are defined by the simple functions.*

Proposition 57. *Let $\mu : \sigma \longrightarrow [0, \infty]$ be a measure. Let $F \in \sigma$. Then for any measurable function f,*

$$\int_F f \, d\mu = \int f \, d\mu|_F$$

Proof. Follows from the way integrals over measurable subsets are defined. □

Proposition 58. *Let μ be a measure, and f an integrable function. Then*

$$\Phi(E) := \int_E f \, d\mu$$

is a measure.

Proof. Can be found in [8]. □

6. FUNCTIONS ON THE REAL LINE

L^2 is a Hilbert space of functions, and most quantum mechanical problems live in L^2. However, we can make spaces of functions which have considerably less properties:

Proposition 59. *The set of functions* $\Omega = \left\{f : \mathbb{R} \longrightarrow \mathbb{C}\right\}$*, equipped with scalar multiplication operation* $(z \cdot f)(x) = z \cdot (f(x))$ *and addition operation* $(f + g)(x) = f(x) + g(x)$ *is a vector space over* $\mathbb{C}$*.*

Notice that Ω here is a very big set. It contains all the real valued functions. It contains the square-integrable functions. It contains literally every possible function you could come up with, subject to the restriction that it takes in real numbers and outputs complex numbers.

The reason it contains so many functions is that we've said so little about the functions it contains. As a result of this, it immediately becomes very difficult to prove anything nontrivial about Ω. The theory of linear operators on Ω would not be particularly fruitful. Intricate structure of a space arises from cleverly chosen requirements for elementhood. If I give you a vector space V and ask you to prove something about it, you're going to have a considerably more difficult time than if I told you it was a normed vector space, because there's no way of establishing any sort of convergence for sequences or operators. If I give you a Banach space, you're going to have an even easier time establishing convergence of sequences, as only the Cauchy requirement is necessary. With a Hilbert space, you may make use of the inner product while proving a result. The more you say about a collection, the less things you have in your collection, so more structure arises.

Moving back down to L^2, we will see that in many cases we already have many pathological functions, and so very often L^2 is not worked with directly. Instead, we very often talk about the Schwartz space, a subspace of L^2.

While constructing these spaces, we often start from the bottom up. The atoms of functional analysis are the aptly named simple functions.

Definition 77. *A function* $s : M \longrightarrow N$ *is said to be simple if it has a finite range* $\operatorname{ran} s = \{s_1, ..., s_n\}$ *The set of measurable simple functions* $s : M \longrightarrow N$ *by* $S(M.N)$*.*

Proposition 60. $S(\mathbb{R}, \mathbb{C})$ *is a vector space over* $\mathbb{C}$*, equipped with the addition operation* $(s + t)(x) := s(x) + t(x)$*,* $(ks)(x) = ks(x)$*.*

Proposition 61. *$S(\mathbb{R},\mathbb{C})$ is a normed vector space, when equipped with the infinity norm*

$$\|s\|_\infty = \sup|s(x)|$$

Proof. First, we note that for any simple s, since the range is finite the $\sup|s(x)|$ is always just the $max|s(x)|$.

(i) $\|s\| \geq 0$

Proof: Each $|s(x)|$ is a non-negative number so the same is true in the max.

(ii) $\|s\| = 0 \implies s = 0$

Proof: If the $max|s(x)| = 0$, then since $|s(x)| \leq max|s(x)|$ for each x and $|s(x)| \geq 0$ for each x we have that $|s(x)| = 0$ for each x so that $s(x) = 0$ for each x.

(iii) $\|\lambda s\| = |\lambda|\|s\|$

Proof: $\|\lambda s\| = max|\lambda s(x)| = max|\lambda||s(x)| = |\lambda|max|s(x)| = |\lambda|\|s\|$.

(iv) $\|s+t\| \leq \|s\| + \|t\|$

Proof: $|s(x)| \leq \|s\|$, and $|t(x)| \leq \|t\|$. So

$$|s(x)+t(x)| \leq |s(x)| + |t(x)| \leq \|s\| + \|t\|$$

so that in taking the max we obtain the desired result.

□

Definition 78. *The set of bounded measurable functions is defined by:*

$$B(\mathbb{R},\mathbb{C}) := \left\{ f : \mathbb{R} \longrightarrow \mathbb{C} \middle| \|f\|_\infty < \infty \text{ , } f \text{ is measurable} \right\}$$

Proposition 62. *$B(\mathbb{R},\mathbb{C})$ is a Banach space, equipped with the usual addition, scalar multiplication, and infinity norm.*

Proof. The vector space proof is left to the reader. The norm follows the same proof with sup facts. (might fill in later) Proof of Banach: Let $f_n \in B$ be Cauchy. Let $\epsilon > 0$ Since f_n is Cauchy $\exists N$: $\|f_n - f_m\| < \epsilon$ whenever $n > N$. But, for each x, $|(f_n - f_m)(x)| \leq \|f_n - f_m\|$. Therefore $\forall x, f_n(x)$ is a Cauchy sequence in $\mathbb{C}$. Since $\mathbb{C}$ is complete, $f_n(x)$ converges for each x. Define the function $f(x) = \lim f_n(x)$. We wish to show that $f \in B$ and $f = \lim f_n$ according to the infinity norm. Choose $\epsilon = 1$. Then $\exists N$:

$$\|f_n - f_m\| < 1 \text{ whenever } n, m > N$$

Then for all x,

$$|f_n(x) - f_m(x)| < 1 \text{ whenever } n, m > N$$

Now since the limit exists pointwise, we take the limit on m:

$$|f_n(x) - f(x)| \leq 1 \text{ whenever } n > N$$

We choose $n = N + 1$ to make the above inequality true. We then have, by the reverse triangle inequality:

$$|f(x)| - |f_{N+1}(x)| \leq 1$$

Rearranging and taking the sup on both sides we have

$$\sup |f(x)| \leq 1 + \|f_{N+1}\|$$

But since f_{N+1} is bounded (also need measurability) we have $f \in B$. The same $\frac{\epsilon}{4}$ trick works here as it did for the bounded linear maps to verify that $f = \lim f_n$. □

The proof of the following proposition was developed independently.

Proposition 63. *$S(\mathbb{R}, \mathbb{R})$ is dense in $B(\mathbb{R}, \mathbb{R})$.*

Proof. We need to show that for each $f \in B$ we can find a sequence $s_n \in S$ such that $\lim s_n = f$ with respect to the infinity norm. Of course, in the context of the infinity norm, this means we have to make $\sup |f(x) - s_n(x)| < \epsilon$ for $n > N$, which means the function s_n has to converge pointwise to f, all using the same N. To do this, consider an f. Since $f \in B$, it must be true that $|f(x)| \leq \|f\|$. Therefore let $N = \inf f(x)$ and $M = \sup f(x)$. Therefore, $\operatorname{ran} f \subseteq [N, M]$. For the simple function s_n, we break up the domain into

$$A_i = f^{-1}\Big[N + (i-1)\frac{M-N}{n}, N + (i)\frac{M-N}{n}\Big) \text{ for } i = 1, ..., n$$

and define $s_n(x) = N + (i-1)\frac{M-N}{n}$ for $x \in A_i$. (Note to self: Missing the very top here. Should have closed bracket.) Since f is measurable, the preimage of each interval is measurable. This construction looks miles more complicated than it actually is. The simple idea is, break up length of the range into n little pieces, then break up the domain into the things that get mapped to each of those pieces. Then define a simple function to be mapped to the bottom of each of those pieces in the range.

Now, let $x \in \mathbb{R}$. Then for any n, $\exists j$ s.t. $1 \leq j \leq n$ and $f(x) \in A_j$. Observe that the inequality

$$\|f(x) - s_n(x)\| \leq \frac{1}{n}$$

follows from the fact that $f(A_j)$ has length $\frac{1}{n}$. Separately, observe that $\frac{1}{n} < \frac{1}{N}$ when $n > N$. At last, let $\epsilon > 0$. Let $N = \frac{2}{\epsilon}$. Then for any x,

$$\|f(x) - s_n(x)\| \leq \|f(x) - s_N(x)\| \leq \frac{1}{N} = \frac{\epsilon}{2} \text{ whenever } n > N$$

Since our choice of N satisfies the desired inequality for all x, we take the sup on both sides and get that

$$\|f - s_n\| \leq \frac{\epsilon}{2} < \epsilon \,\text{whenever}\, n > N$$

Therefore $f = \lim s_n$. But this is true for any $f \in B(\mathbb{R}, \mathbb{R})$. Therefore $S(\mathbb{R}, \mathbb{R})$ is dense in $B(\mathbb{R}, \mathbb{R})$. □

Proposition 64. *$S(\mathbb{R}, \mathbb{C})$ is dense in $B(\mathbb{R}, \mathbb{C})$.*

Proof. Uses the same trick as before, except instead of s_n mapping the line into n intervals, it maps the line into an nxn grid of the plane with side length $2\|f\|$. □

Definition 79. *The set of square integrable functions (with respect to the measure μ) from $\mathbb{R}$ to $\mathbb{C}$ is defined by*

$$\mathcal{L}^2(\mathbb{R}, \mu) := \left\{ f : \mathbb{R} \longrightarrow \mathbb{C} \middle| \int |f|^2 \, d\mu < \infty \right\}$$

Remark 14. *If μ is the Lebesgue measure, $\mathcal{L}^2(\mathbb{R}, \mu)$ is just denoted $\mathcal{L}^2(\mathbb{R})$, or sometimes just $\mathcal{L}^2$.*

This set is, unfortunately, not a Hilbert space. For consider the function $\xi(x) = 0$ if $x \neq 0$ but $\xi(x) = 1$ if $x = 0$. This function is not zero, but its square integral is 0, so we will run into trouble with the norm axioms if we're not careful. To avoid this, we define an equivalence relation.

Definition 80. *We define an equivalence relation $\equiv$ for $\mathcal{L}^2(\mathbb{R})$ in the following way: $f \equiv g$ if $\exists M$ with $\mu(M) = 0$ such that*

$$f(x) = g(x) \, \forall x \in \mathbb{R} \backslash M$$

And we denote the set of elements that are equivalent to some f by $[f]$.

Proposition 65. *If $f \equiv g$, then $\int f \, d\mu = \int g \, d\mu$ and $\int |f|^2 \, d\mu = \int |g|^2 \, d\mu$.*

Proof. Let M be the set where f and g perhaps differ. Then

$$\int f \, d\mu = \int_M f \, d\mu + \int_{M^c} f \, d\mu = \int_{M^c} g \, d\mu = \int_M g \, d\mu + \int_{M^c} g \, d\mu = \int g \, d\mu$$

□

Definition 81. *Let $\mathcal{L}^2$ be defined as above. The set of equivalence classes of square integrable functions is defined by*

$$L^2(\mathbb{R}) := \{ [f] \, | f \in \mathcal{L}^2 \}$$

Now, when dealing with equivalence classes of functions, we should technically be careful to define something like $[f] + [g] =: [a + b]$, where we are allowed to pick any $a \in [f]$ and $b \in [g]$, and then check that such operations are well-defined, but in practice, one never really has to be worried about this, and as long as we understand that if we somehow end up with a function like ξ, we just talk about it as the 0 function. In the following definition we omit mention of equivalence classes.

Theorem 6. *$L^2(\mathbb{R})$ is a Hilbert space when equipped with function addition and scalar multiplication and the inner product*

$$\langle f, g \rangle := \int f^* g \, d\mu$$

Proof. Can be found in [9]. □

As we have previously discussed, L^2 is sometimes hard to work with concretely. The following space is immensely useful for quantum mechanical operators:

Definition 82. *The Schwartz space of $\mathbb{R}$ is defined by*

$$\mathcal{S}(\mathbb{R}) := \left\{ f \in C^\infty \middle| \sup |x^\alpha D^\beta f(x)| < \infty \right\}$$

Proposition 66. *$\mathcal{S}(\mathbb{R})$ is dense in $L^2(\mathbb{R})$.*

Proof. Can be found in [9]. □

Proposition 67. *The Schwartz space is invariant under the position and momentum operators.*

Proof. Can be found in [9]. □

7. SPECTRAL MEASURES

The time evolution of a quantum system is generated by the Hamiltonian operator of the system. The time evolution operator can be written down by the equation

$$U(t) := \exp \frac{-iHt}{\hbar}$$

This should remind the reader of Schrödinger's equation. For if we cheat, we may observe that the equation

$$i\hbar D_t \Psi(t) = H\Psi(t)$$

is kind of like a first order differential equation in t, so if we solve for $\Psi(t)$ we obtain

$$\Psi(t) = e^{\frac{-iHt}{\hbar}} \Psi(0)$$

However, we have no means to define something like $e^{operator}$ yet. The tool that is used to deal with this is the spectral theorem. We want to write

$$A = \int \lambda \, P(d\lambda)$$

So that in effect A acting on a vector ψ acts like projecting the state into a state corresponding to measuring the quantity being between λ and $\lambda + d\lambda$ and then multiplying that by λ, and then summing over all possible measurement values on the real line. First, however, we must define the projection valued measure:

Definition 83. *A function $P : \sigma(\mathcal{O}_{\mathbb{R}}) \longrightarrow \mathcal{L}(H, H)$ is said to be a projection valued measure if it satisfies the three axioms:*

(i) $P(E)$ *is a projector for each* $E \in \sigma$

(ii) $P(\mathbb{R}) = id_H$

(iii) $P(\cup E_i)\psi = \sum \Big(P(E_i)\psi \Big)$ *for any* $\psi \in H$*, and for a sequence of disjoint measurable sets* E_i

Another tool useful for the construction of this PVM is a complex valued measure:

Definition 84. *A function $\mu : \sigma(\mathcal{O}_{\mathbb{R}}) \longrightarrow \mathbb{C}$ is said to be a complex valued measure if it satisfies two axioms:*

(i) $\mu(\emptyset) = 0$

(ii) $\mu(\cup E_i) = \sum \mu(E_i)$ *for a sequence of disjoint measurable sets* E_i

Remark 15. *The axioms for real valued measures are the same for complex valued measures. The only difference is the codomain.*

First, we observe some properties of a PVM. Notice that they mirror the properties of a real valued measure.

Proposition 68. *A PVM has the following properties:*

(i) $P(\emptyset) = 0_H$

(ii) $P(E^c) = 1_H - P(E)$

(iii) $P(E \cup F) + P(E \cap F) = P(E) + P(F)$

(iv) $P(E \cap F) = P(E)\,P(F)$

Proof. Can be found in [9] □

Proposition 69. *Let P be a PVM. For any $\psi, \phi \in H$, the function $\mu_{\psi,\phi} : \sigma(\mathcal{O}_{\mathbb{R}}) \longrightarrow \mathbb{C}$ determined by the equation:*

$$\mu_{\psi,\phi}(E) = \langle \psi, P(E)\phi \rangle$$

is a complex valued measure.

Proof. $P(\emptyset) = 0$ so $\mu_{\psi,\phi} = 0$, and the countable additivity follows trivially from the additivity for P. □

Proposition 70. *Let P be a PVM. For any $\psi \in H$, the function $\mu_\psi : \sigma(\mathcal{O}_{\mathbb{R}}) \longrightarrow \mathbb{R}$ determined by the equation:*

$$\mu_\psi(E) = \mu_{\psi,\psi}(E)$$

is a real valued measure.

Proof. The fact that μ_ψ is a measure follows immediately from the fact that $\mu_{\psi,\psi}$ is a measure. It remains to show that it is real valued. Let $E \in \sigma$. Then

$$\mu_\psi(E) = \mu_{\psi,\psi}(E) = \langle \psi, P(E)\psi \rangle = \langle P(E)^*\psi, \psi \rangle = \langle P(E)\psi, \psi \rangle = \overline{\langle \psi, P(E)\psi \rangle} = \overline{\mu_{\psi,\psi}(E)} = \overline{\mu_\psi(E)} \quad (7.1)$$

□

With these tools, we begin our construction of integrals over a PVM.

Definition 85. *For simple measurable functions $s : \mathbb{R} \longrightarrow \mathbb{C}$, we define the integral of s over P according to the equation:*

$$\int s\, dP := \sum_{i=1}^{n} s_i P(A_i)$$

Remark 16. *This is exactly analogous to the ordinary Lebesgue integral for wave functions.*

Proposition 71. *For any simple function $s \in S(\mathbb{R}, \mathbb{C})$ and any $\psi \in H$, the relation*

$$\| \int s\, dP\, \psi \|^2 = \int |s|^2\, d\mu_\psi$$

holds.

Proof.

$$\| \int s\, dP\, \psi \|^2 = \langle \sum_{j=1}^{n} s_j P(A_j)\psi, \sum_{i=1}^{n} s_i P(A_i)\psi \rangle =$$

$$\langle \psi, \sum_{j=1}^{n} \sum_{i=1}^{n} \overline{s_j} s_i P(A_j)^* P(A_i)\psi \rangle = \langle \psi, \sum_{i=1}^{n} |s_i|^2 P(A_i)\psi =$$

$$\sum_{i=1}^{n} |s_i|^2 \langle \psi, P(A_i)\psi \rangle = \int |s|^2\, d\mu_\psi \quad (7.2)$$

□

Proposition 72. *The operator $\int dP : S(\mathbb{R}, \mathbb{C}) \longrightarrow \mathcal{L}(H)$ defined by the equation:*

$$(\int dP)(s) := \int s\, dP$$

is a bounded linear map with operator norm $\| \int dP \| = 1$.

Proof. Before starting the proof, we make careful note of the fact that we're defining a linear map that maps simple functions into more linear maps on a Hilbert space, so as not to lose track of which linearity we are talking about where.

Now, let $s \in S(\mathbb{R}, \mathbb{C})$. Then $\int s\, dP$ is a finite sum of constant multiples of projectors, which are bounded linear maps, so that the resulting map is again a bounded linear map from H to H. Now, we calculate linearity. Let s and t be simple functions with ranges $s_1, ..., s_n$, and $t_1, ..., t_m$ with preimages labeled

A_i and B_j respectively. Now we calculate

$$\int s\,dP + \int t\,dP = \sum_{i=1}^{n} s_i P(A_i) + \sum_{j=1}^{m} t_j P(B_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} s_i P(A_i \cap B_j) + \sum_{j=1}^{m}\sum_{i=1}^{n} t_j P(B_j \cap A_i) = \sum_{i=1}^{n}\sum_{j=1}^{m}(s_i + t_j)P(A_i \cap B_j) = \int s + t\,dP \quad (7.3)$$

Therefore the map $\int dP$ is linear. To show boundedness, let $\psi \in H$, and let s be any simple function. The by Proposition 71, we have

$$\|\int s\,dP\,\psi\|^2 = \int |s|^2\,d\mu_\psi \leq \int \|s\|^2\,d\mu_\psi = \|s\|^2 \int d\mu_\psi = \|s\|^2\|\psi\|^2 \quad (7.4)$$

And equality holds when s is a constant function, because then each $|s(x)|^2 = \|s\|^2$. Therefore, for any $s \in S(\mathbb{R}, \mathbb{C})$ and any $\psi \in H$.

$$\|\int s\,dP\,\psi\| \leq \|s\|\|\psi\|$$

And equality holds when s is constant.

Let $R > 0$. Consider the family of simple functions S_R which have $\|s\| = R$. Let s_R be the constant function $s_R(x) = R$. Then for any $s \in S_R$,

$$\|\int s\,dP\,\psi\| \leq \|s\|\|\psi\| = \|\int s_R\,dP\,\psi\|$$

Now consider any simple function $s \neq 0$. Clearly it belongs to S_R for some $R > 0$, so that when looking at the operator norm, we only need to take the sup over positive constant simple functions.

$$\|\int dP\| = \sup_{R>0} \frac{\|\int s_R\,dP\|}{\|s_R\|} = \sup_{R>0}\sup_{\psi\neq 0} \frac{\|\int s_R dP\psi\|}{\|s_R\|\|\psi\|} = \sup_{R>0}\sup_{\psi\neq 0} \frac{R\|\psi\|}{R\|\psi\|} = \sup\sup 1 = 1 \quad (7.5)$$

□

Corollary 10. *Since $S(\mathbb{R}, \mathbb{C})$ is dense in $B(\mathbb{R}, \mathbb{C})$, we have by the BLT theorem that the bounded linear transformation $\int dP$ can be extended to all $B(\mathbb{R}, \mathbb{C})$.*

Remark 17. *For any bounded measurable f, the integral of f over P is defined by the equation*

$$\int f\,dP := \lim \int s_n\,dP$$

for $f = \lim s_n$, where the lim *here uses the norm of $B(\mathbb{R}, \mathbb{C})$.*

We now move to general measurable functions which are not necessarily bounded. We note that these integrals will encompass unbounded operators, and it is with general measurable functions where the integral can only be defined on a dense subset.

Definition 86. *For any measurable function $f : \mathbb{R} \longrightarrow \mathbb{C}$, we first define the set*

$$D_{\int f\,dP} := \Big\{ \psi \in H \Big| \int |f|^2 \, d\mu_\psi < \infty \Big\}$$

Before defining the integral, we approximate f by a sequence of bounded measurable functions:

$$f_n = f \circ \chi_{A_n}$$

where $A_n = \Big\{ x \in \mathbb{R} \;\Big|\; |f(x)| < n \Big\}$ and then we define the operator $\int f\,dP : D_{\int f\,dP} \longrightarrow H$ by the equation:

$$\Big(\int f\,dP \Big)\psi := \lim \Big(\int f_n \, dP\psi \Big)$$

The proof of the following proposition was developed independently, but is similar to the proof of [8] using resolutions of the identity. I believe this proof to be the cleanest, especially for the spectral theorem defined in terms of PVMs.

Proposition 73. *The operator defined in Definition 86 is a densely defined linear operator, with maximal domain.*

Proof. First, we show well-definition of the operator for any ψ in the domain. In the $L_2(\mu_\psi)$ norm, we have

$$\|f_n - f\|^2 = \int |f_n - f|^2 \, d\mu_\psi = \int_{F_n^c} |f|^2 \, d\mu_\psi = \Phi(F_n^c)$$

Where Φ is a measure defined in the obvious way. Observe that

$$\bigcup_{n=1}^{\infty} F_n = \mathbb{R}$$

so that

$$\bigcap_{n=1}^{\infty} F_n^c = \emptyset$$

We have therefore that

$$0 = \Phi(\emptyset) = \Phi(\bigcap_{n=1}^{\infty} F_n^c) = \lim \Phi(F_n) = \lim \|f_n - f\|^2$$

So that $\lim f_n = f$ in $L_2(\mu_\psi)$. Since $L_2(\mu_\psi)$ is a Hilbert space, f_n is Cauchy in $L_2(\mu_\psi)$ for free. Moving back to the operators, we can use this:

$$\Big\| \int f_n \, dP\psi - \int f_m \, dP\psi \Big\|^2 = \Big\| \int f_n - f_m \, dP\psi \Big\|^2 = \int |f_n - f_m|^2 \, d\mu_\psi = \|f_n - f_m\|_{L_2(\mu_\psi)} \quad (7.6)$$

So that the $\int f_n \, dP\psi$ is Cauchy in the regular L_2. Since L_2 is complete, the $\int f_n \, dP\psi$ converges. Therefore $\int f \, dP$ is well defined for ψ in its domain.

We show now that the domain is dense. Let $\psi \in H$. Define $\psi_n = P(F_n)\psi$. Observe that

$$\mu_{\psi_n}(E) = \langle \psi_n, P(E)\psi_n \rangle = \langle P(F_n)\psi, P(E)P(F_n)\psi \rangle = \langle \psi, P(F_n)P(E \cap F_n)\psi \rangle = \langle \psi, P(E \cap F_n)\psi \rangle = \mu_\psi(E \cap F_n) \quad (7.7)$$

so that by proposition, any integral with respect to μ_{ψ_n} amounts to an integral over F_n with respect to μ_ψ. That is,

$$\int |f|^2 \, d\mu_{\psi_n} = \int_{F_n} |f|^2 \, d\mu_\psi = \int |f_n|^2 \, d\mu_\psi = \| \int f_n \, dP\psi \|^2 \leq \|f_n\|^2 \|\psi\|^2 \leq n^2 \|\psi\|^2 < \infty \quad (7.8)$$

so that each $\psi_n \in D_{\int f \, dP}$. In fact, these ψ_n's approach ψ.

$$\psi = I\psi = P(\bigcup_{n=1}^{\infty} F_n)\psi = \lim P(F_n)\psi = \lim \psi_n$$

So that the domain is dense in H.

Linearity is left to the reader, and follows from the properties of limits.

To prove that the domain is maximal, observe that if

$$\lim \int f_n \, dP\psi =: \phi$$

exists, then the quantity

$$\lim \| \int f_n \, dP\psi \|^2 \leq (1 + \|\phi\|^2) < \infty$$

But from measure theory we have

$$\int |f|^2 \, d\mu_\psi = \Phi(\mathbb{R}) = \lim \Phi(F_n) = \lim \int |f_n|^2 \, d\mu_\psi = \lim \| \int f_n \, dP\psi \|^2$$

so that f is μ_ψ square-integrable, i.e. the ψ in the domain coincide with the ψ for which the limit exists. □

Proposition 74. *For any measurable function $f : \mathbb{R} \longrightarrow \mathbb{C}$, the following properties hold:*

(i) $\Big(\int f\,dP\Big)^* = \int \overline{f}\,dP$

(ii) $\int(\alpha f + g)\,dP \supseteq \alpha \int f dP + \int g\,dP$

(iii) $\int f \cdot g\,dP \supseteq \int f\,dP \circ \int g\,dP$

In the case that f and g are bounded, equality holds for (ii) and (iii).

Proof. We shall only prove (i); the proof of the others can be found in [9]. In order to prove the general case, we build up again from the case that the function is simple and measurable.

$$\Big(\int s\ dP\Big)^* = \Big(\sum_{i=1}^{n} s_i P(A_i)\Big)^* = \sum_{i=1}^{n} \overline{s_i} P(A_i)^* = \sum_{i=1}^{n} \overline{s_i} P(A_i) = \int \overline{s}\,dP \quad (7.9)$$

Now suppose f is bounded and measurable, with $f = \lim s_n$. Then $\overline{f} = \lim \overline{s_n}$. Now, let $\psi, \phi \in H$. Then

$$\begin{aligned}\langle\Big(\int f\,dP\Big)^*\psi, \phi\rangle = \langle\psi, \int f\,dP\phi\rangle = \langle\psi, \lim\Big(\int s_n\,dP\Big)\phi = \\ \langle\psi, \lim\Big(\int s_n\,dP\phi\Big)\rangle = \lim\langle\int \overline{s_n}\,dP\psi, \phi\rangle = \langle\lim\Big(\int \overline{s_n}\,dP\psi\Big), \phi\rangle = \\ \langle\lim\Big(\int \overline{s_n}\,dP\Big)\psi, \phi\rangle = \langle\int \overline{f}\,dP\psi, \phi\rangle \quad (7.10)\end{aligned}$$

But this holds for all $\psi, \phi \in H$, so the operators at the beginning and end must be equal. Now we move to measurable f. Define f_n as usual. Since $|f|^2 = |\overline{f}|^2$, clearly $D_{\int f\,dP} = D_{\int \overline{f}\,dP}$. Now let $\psi \in D_{\left(\int f\,dP\right)^*}$ and $\phi \in D_{\int f\,dP}$ Then

$$\begin{aligned}\langle\Big(\int f\,dP\Big)^*\psi, \phi\rangle = \langle\psi, \Big(\int f\,dP\Big)\phi\rangle = \langle\psi, \lim\Big(\int f_n\,dP\phi\Big)\rangle = \\ \lim\langle\psi, \int f_n\,dP\phi\rangle = \lim\langle\Big(\int f_n\,dP\Big)^*\psi, \phi\rangle = \\ \lim\langle\int \overline{f_n}\,dP\psi, \phi\rangle = \langle\lim\Big(\int \overline{f_n}\,dP\psi\Big), \phi\rangle = \langle\int \overline{f}\,dP\psi, \phi\rangle \quad (7.11)\end{aligned}$$

□

Now that we have a complete theory of the integration of functions with respect to PVMS, we state the following proposition:

Proposition 75. *Given a PVM P, the operator $A : D_A \longrightarrow H$ defined by*

$$A = \int id_{\mathbb{R}}\, dP$$

is a self-adjoint operator.

Proof. $\overline{id_{\mathbb{R}}} = id_{\mathbb{R}}$. □

Theorem 7. *(Spectral theorem) For every self-adjoint operator $A : D_A \longrightarrow H$, there exists a unique PVM P such that:*

$$A = \int id_{\mathbb{R}}\, dP$$

Proof. Can be found in **??** □

We detail the construction of P given A to make Theorem 7 true. In effect, instead of starting with P and building the μ measures, we start with A, and work backwards to build P.

Proposition 76. *Given a self-adjoint operator $A : D_A \longrightarrow \mathcal{H}$, then the measure $\mu_\psi : \sigma(\mathcal{O}_{\mathbb{R}}) \longrightarrow \mathbb{R}$ can be solved for according to an integral:*

$$\mu_\psi(-\infty, \lambda] := \frac{1}{\pi} \lim_{\delta \to 0^+} \lim_{\epsilon \to 0^+} \int_{-\infty}^{\lambda+\delta} Im\, \langle \psi, (A - t - i\epsilon)^{-1} \psi \rangle dt$$

From this, the measure $\mu_{\psi,\phi} : \sigma(\mathcal{O}_{\mathbb{R}}) \longrightarrow \mathbb{C}$ can be easily rewritten in terms of the polarization identity:

$$\mu_{\psi,\phi}(E) := \frac{1}{4}\Big[\mu_{\psi+\phi} - \mu_{\psi-\phi} + i\mu_{\psi-i\phi} - i\mu_{\psi+i\phi}\Big](E)$$

From this, the PVM can be reconstructed $P : \sigma(\mathcal{O}_{\mathbb{R}}) \longrightarrow \mathcal{L}(H)$ according to the equation:

$$\langle \psi, P(E)\phi \rangle := \mu_{\psi,\phi}(E)$$

Proof. The proof can be found in **??**. □

But what does a PVM look like? It helps to look at the μ_ψ's instead to get a feel for how the PVM sees everything up to a number λ.

Proposition 77. *The PVM for the map $I : H \longrightarrow H$ is given by $P(E) = I$ if $1 \in E$ and $P(E) = 0$ otherwise.*

Proof. We start by computing μ_ψ for a given ψ, as per Proposition 76:

$$\mu_\psi(-\infty, \lambda] = \frac{1}{\pi} \lim_{\delta \to 0^+} \lim_{\epsilon \to 0^+} \int_{-\infty}^{\lambda+\delta} \mathrm{Im}\langle \psi, [I - (t + i\epsilon)I]^{-1} \psi \rangle dt$$

But observe that $[I - (t + i\epsilon)I]^{-1}$ is just $(1 - t + i\epsilon)^{-1} I$. We can thus pull the real number $\langle \psi, \psi \rangle$ out from under the "Im", the integral and all the limits to

get it out of the way. In addition, for any $z \neq 0$, it is true that $\frac{1}{z} = \frac{\overline{z}}{|z|^2}$, so we have

$$\begin{aligned}\mu_\psi(-\infty, \lambda] &= \frac{1}{\pi}\|\psi\|^2 \lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\int_{-\infty}^{\lambda+\delta} \frac{\epsilon}{(1-t)^2+\epsilon^2}dt \\ &= \frac{1}{\pi}\|\psi\|^2 \lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\Big[-\arctan\Big(\frac{1-t}{\epsilon}\Big)\Big]_{-\infty}^{\lambda+\delta} \\ &= \frac{1}{\pi}\|\psi\|^2 \lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\Big[\arctan\Big(\frac{1-t}{\epsilon}\Big)\Big]_{\lambda+\delta}^{-\infty} \\ &= \frac{1}{\pi}\|\psi\|^2 \lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\frac{\pi}{2} - \arctan\Big(\frac{1-(\lambda+\delta)}{\epsilon}\Big) \quad (7.12)\end{aligned}$$

Consider the case $\lambda < 1$. Then we can find a small enough δ so that $\lambda + \delta < 1$ still, i.e. $1 - (\lambda + \delta) > 0$. Then clearly as ϵ gets smaller and smaller from the right, the quantity arctan(...) in the final equation above gets closer and closer to $\frac{\pi}{2}$, so the whole thing just becomes zero. Then taking the limit on delta, we see that

$$\mu_\psi(-\infty, \lambda] = 0 \quad \text{if} \quad \lambda < 0$$

if $\lambda \geq 1$, however, then for any positive δ $\lambda + \delta > 1$, so that $1 - (\lambda + \delta) < 0$. In this case, as ϵ gets closer to 0, the *arctan*(...) term becomes $-\frac{\pi}{2}$. Taking the limit on δ from the right preserves the positivity of the term inside of arctan so that:

$$\mu_\psi(-\infty, \lambda] = \|\psi\|^2 \quad \text{if} \quad \lambda \geq 0$$

Then polarize. It then follows that, whenever $E \in \sigma(\mathcal{O}_\mathbb{R})$ contains the number 1,

$$\langle \phi, P(E)\psi\rangle = \langle \phi, \psi\rangle$$

But subtracting the right side from the left one sees that $P(E)$ is in this case the identity. If we now suppose E does not contain 1, then for each ψ it holds that

$$0 = \langle \psi, P(E)\psi\rangle = \langle \psi, P(E)^2\psi\rangle = \langle P(E)\psi, P(E)\psi\rangle = \|P(E)\psi\|^2$$

so that each $P(E)\psi$ must be the zero vector. □

But what is special about the point 1? Well, it's the only number of the spectrum of I. This is the fundamental characteristic of these PVMs.

While we have not yet fully characterized the position and momentum operators, the reader may assume the position operators have the ordinary properties that are asserted in most textbooks, *i.e.*$Q\psi(x) = x\psi(x)$.

7.1 The Position Operator

What is the PVM for the position operator Q: $\mathcal{S}(R) \longrightarrow \mathcal{S}(R)$? Well, let ψ in $\mathcal{S}$. Then $(Q\psi)(x) = x\psi(x)$. Therefore for any complex number μ, we have

$(Q-\mu)\psi(x) = (x-\mu)\psi(x)$. If $\mu \in \mathbb{C}\backslash\mathbb{R}$, then it is in the resolvent so it has an inverse, given by

$$((Q-\mu)^{-1}\psi)(x) = \frac{1}{x-\mu}\psi(x)$$

Plugging this into the μ_ψ formula, we obtain

$$\begin{aligned}
\mu_\psi^Q(-\infty,\lambda] &= \frac{1}{\pi}\lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\int_{-\infty}^{\lambda+\delta}\operatorname{Im}\langle\psi,(Q-(t+i\epsilon))^{-1}\psi\rangle dt \\
&= \frac{1}{\pi}\lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\int_{-\infty}^{\lambda+\delta}\operatorname{Im}\int\psi^*(x)[(Q-(t+i\epsilon))^{-1}\psi](x)\,\mu(dx)\,dt \\
&= \frac{1}{\pi}\lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\int_{-\infty}^{\lambda+\delta}\operatorname{Im}\int\frac{1}{x-(t+i\epsilon)}|\psi(x)|^2\,\mu(dx)\,dt \\
&= \frac{1}{\pi}\lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\int_{-\infty}^{\lambda+\delta}\int\operatorname{Im}\frac{1}{x-(t+i\epsilon)}|\psi(x)|^2\,\mu(dx)\,dt \\
&= -\frac{1}{\pi}\lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\int_{-\infty}^{\lambda+\delta}\int\frac{\epsilon}{(x-t)^2+\epsilon^2}|\psi(x)|^2\,\mu(dx)\,dt \qquad (7.13)
\end{aligned}$$

Then we exchange the integral signs, and evaluate:

$$\begin{aligned}
&= -\frac{1}{\pi}\lim_{\delta\to 0^+}\lim_{\epsilon\to 0^+}\int\Big[\arctan\Big(\frac{x-t}{\epsilon}\Big)\Big]_{-\infty}^{\lambda+\delta}|\psi(x)|^2\,\mu(dx) \\
&= -\frac{1}{\pi}\lim_{\delta\to 0^+}\int\lim_{\epsilon\to 0^+}\Big[\arctan\Big(\frac{x-t}{\epsilon}\Big)\Big]_{-\infty}^{\lambda+\delta}|\psi(x)|^2\,\mu(dx) \qquad (7.14)
\end{aligned}$$

But using similar techniques as for the identity, we obtain

$$= \lim_{\delta\to 0^+}\int_{(-\infty,\lambda+\delta)}|\psi|^2 d\mu$$

But using the fact that this integral is a measure over the set it integrates over, we can take the limit to get inclusion, so for the case of a general set, we have

$$\mu_\psi^Q(E) = \int_E |\psi(x)|^2\,dx$$

Of *course* it's the integral of $|\psi|^2$, we're talking about the probability of measuring the particle in the set E while the system is in the state ψ. That is,

$$Prob(E) = Tr(P^Q(E)\rho) = \langle\psi, P^Q(E)\psi\rangle = \int_E |\psi(x)|^2\,dx$$

7.2 The Momentum Operator

The momentum operator P and the position operator Q are the most important observables. Any observable found in an equation is usually a function of P and

Q. What is the PVM for the momentum operator P? It turns out, we've already done all the hard work when we did the calculation for Q, as they are related by the Fourier transform. For notational convenience, we denote the momentum operator as M for the time being.

$$M = UQU^{-1}$$

So we observe the quantity

$$(M - zI)^{-1} = (UQU^{-1} - zI)^{-1}$$

Using the fact that I commutes with any operator and that $UU^{-1} = I$, we have that the above

$$= (UQU^{-1} - U(zI)U^{-1})^{-1} = (U(Q - zI)U^{-1})^{-1} = U^{-1}(Q - zI)^{-1}U$$

so that for any $\psi \in S$ we have

$$\langle \psi, (M - z)^{-1}\psi \rangle = \langle \psi, U^{-1}(Q - z)^{-1}U\psi \rangle = \langle U\psi, (Q - z)^{-1}U\psi \rangle$$

But this is exactly the quantity appearing in the formula for μ_ψ^M. Therefore given $\psi \in S$ and $E \in \sigma$ we have

$$\mu_\psi^M(E) = \mu_{U\psi}^Q(E)$$

and

$$P^M(E) = U^{-1}P^Q(E)U$$

But restated, this is just

$$\langle \psi, P^M(E)\psi \rangle = \int_E |\hat{\psi}(p)|^2 dp$$

Where $\hat{\psi}$ is the usual Fourier transform of ψ. This is again of course what we should arrive at: The probability of finding the value of the momentum in the set E is the integral of $\hat{\psi}$ over E in momentum space.

7.3 Multiplication Operators

Now we move to the case of operators defined by the multiplication of a function, i.e. given $g : \mathbb{R} \longrightarrow \mathbb{R}$ and $H = L^2$ we define $G : D_G \longrightarrow H$ defined by $(G\psi)(x) = g(x)\psi(x)$. Then one may verify that for any $z \in \mathbb{C}\backslash\mathbb{R}$

$$((G - z)^{-1}\psi)(x) = \frac{1}{g(x) - z}\psi(x)$$

But this is just the same as for the operator Q with x replaced by $g(x)$. Following the same steps in the case of the position operator, one obtains

$$\mu_\psi(-\infty, \lambda] = ... = -\frac{1}{\pi}\lim_{\delta \to 0^+}\lim_{\epsilon \to 0^+}\int \Big[\arctan\Big(\frac{g(x) - t}{\epsilon}\Big)\Big]_{-\infty}^{\lambda+\delta} dx$$
$$= \int_{g^{-1}(-\infty,\lambda]} |\psi(x)|^2 \, dx \quad (7.15)$$

And by the same method for Q we thus obtain

$$\langle \psi, P^G(E)\psi \rangle = \int_{g^{-1}(E)} |\psi(x)|^2 \, dx$$

Corollary 11. *For the potential energy operator V we have*

$$\langle \psi, P^V(E)\psi \rangle = \int_{V^{-1}(E)} |\psi(x)|^2 \, dx$$

7.4 Operators which are Unitarily Equivalent to Multiplication

The momentum was a special case of an operator which is unitarily equivalent to the the multiplication by the function $f(x) = x$. Moving to the general case, let $G : D_G \longrightarrow H$ be a self-adjoint operator generated by $g : \mathbb{R} \longrightarrow \mathbb{R}$ and B be an operator which is unitarily equivalent to G, i.e. there exists a unitary U such that $B = UGU^{-1}$. Then for any $E \in \sigma(\mathcal{O}_{\mathbb{R}})$,

$$P^B(E) = U^{-1} P^G(E) U$$

or in other words,

$$\langle \psi, P^B(E)\psi \rangle = \int_{g^{-1}(E)} |U\psi(x)|^2 dx$$

7.5 PVMs for Pure Point Spectra

Proposition 78. *If an operator $B : D_B \longrightarrow H$ is self-adjoint and its inverse $B^{-1} : H \longrightarrow D_B$ exists in $\mathcal{L}(H)$, then B^{-1} is self-adjoint.*

Proof.

$$D_{B^{-1*}} = \left\{ \psi \in H \middle| \exists \eta \in H : \forall \alpha \in D_{B^{-1}}, \langle \psi, B^{-1}\alpha \rangle = \langle \eta, \alpha \rangle \right\}$$

Take a $\psi \in H$. $H = \operatorname{ran}(A)$, so $\psi = B\phi$ for some $\phi \in D_B$, and therefore $B^{-1}\psi = \phi$. Therefore

$$\langle \psi, B^{-1}\alpha \rangle = \langle B\phi, B^{-1}\alpha \rangle = \langle \phi, BB^{-1}\alpha \rangle = \langle \phi, \alpha \rangle$$

So $D_{B^{-1*}} = H$ and $B^{-1*}\phi = B^{-1}\phi$. This completes the proof.

□

Lemma 2. *Let $A : D_A \longrightarrow H$ be a densely defined self-adjoint operator with $\sigma(A) \neq \mathbb{R}$. Let $t \in \mathbb{R} \setminus \sigma(A)$. Then for each $\psi \in H$,*

$$\lim_{\epsilon \to 0} (A - t - i\epsilon)^{-1}\psi = (A - t)^{-1}\psi$$

Proof. Define $B = A - t$ for shorthand. For any $\epsilon > 0$, we have

$$\begin{aligned}\|(B-i\epsilon)^{-1}\| &= \sup_{\psi\in H\setminus\{0\}} \sqrt{\frac{\langle (B-i\epsilon)^{-1}\psi, (B-i\epsilon)^{-1}\psi\rangle}{\langle\psi,\psi\rangle}} \\ &= \sup_{\phi\in D_B\setminus\{0\}} \sqrt{\frac{\langle\phi,\phi\rangle}{\langle (B-i\epsilon)\phi, (B-i\epsilon)\phi\rangle}} = \sup_{\phi\in D_B\setminus\{0\}} \sqrt{\frac{\langle\phi,\phi\rangle}{\langle B\phi, B\phi\rangle + \epsilon^2\langle\phi,\phi\rangle}} \\ &\leq \sup_{\phi\in D_B\setminus\{0\}} \sqrt{\frac{\langle\phi,\phi\rangle}{\langle B\phi, B\phi\rangle}} = \sup_{\psi\in H\setminus\{0\}} \sqrt{\frac{\langle B^{-1}\psi, B^{-1}\psi\rangle}{\langle\psi,\psi\rangle}} = \|B^{-1}\| \end{aligned} \tag{7.16}$$

Let $\psi \in H$. Define $\phi(\epsilon) = (B - i\epsilon)^{-1}\psi$. We wish to show then that

$$\lim_{\epsilon\to 0} \phi(\epsilon) = \phi(0)$$

Observe that for each $\epsilon > 0$ we have

$$(B - i\epsilon)(\phi(\epsilon) - \phi(0)) = -i\epsilon\phi(0) = -i\epsilon B^{-1}\psi$$

so that

$$\begin{aligned}&\|\phi(\epsilon) - \phi(0)\| = \|(B-i\epsilon)^{-1}(B-i\epsilon)(\phi(\epsilon)-\phi(0))\| \\ &\leq \|(B-i\epsilon)^{-1}\|\|(B-i\epsilon)(\phi(\epsilon)-\phi(0))\| = \|B^{-1}\|\|-i\epsilon B^{-1}\psi\| \leq \epsilon\|B^{-1}\|^2\|\psi\|^2\end{aligned} \tag{7.17}$$

from which it follows that

$$\lim_{\epsilon\to 0} \phi(\epsilon) = \phi(0)$$

This completes the proof. □

Proposition 79. *If a set E does not contain any of the spectrum of A, then $\forall\psi \in H$,*

$$\lim_{\epsilon\to 0^+} \int_E Im\,\langle\psi, R_A(t+i\epsilon)\psi\rangle dt = 0$$

Proof. Let $t \in E$. Then certainly t is real, so that $A-t$ is a self-adjoint operator. Since $t \in \rho(A)$ by assumption, we also have that $(A-t)^{-1}$ is defined on the whole Hilbert space H. Therefore, $(A-t)^{-1}$ is self-adjoint. Therefore given any $\psi \in H$, the quantity

$$\langle\psi, (A-t)^{-1}\psi\rangle \in \mathbb{R}$$

Since $\sigma(A) \subseteq \mathbb{R}$, we have for free any $\epsilon \geq 0$ that $(A - (t + i\epsilon))^{-1}$ is defined on the whole space, and by Lemma 2 we have that for each ψ

$$\lim_{\epsilon\to 0}(A - (t+i\epsilon))^{-1}\psi = (A-t)^{-1}\psi$$

so that

$$\lim_{\epsilon\to 0}\langle\psi, (A-(t+i\epsilon))^{-1}\psi\rangle = \langle\psi, (A-t)^{-1}\psi\rangle$$

Since the latter quantity is in $\mathbb{R}$, we have

$$\operatorname{Im} \lim_{\epsilon \to 0} \langle \psi, (A - (t + i\epsilon))^{-1} \psi \rangle = 0$$

We can swap the imaginary part with the limit, and integrate the zero function of t over E:

$$\int_E \lim_{\epsilon \to 0} \operatorname{Im} \langle \psi, R_A(t + i\epsilon)\psi \rangle \, dt = 0$$

But since the limit exists, we may pull it out of the integral. This completes the proof. □

We now move to the case where $\lambda \in \sigma_{pp}(A)$ but $\psi \in N(A-\lambda)^\perp = ran(A-\lambda)$. We use the symbol $\Delta := \operatorname{ran}(A - \lambda) \cap D_A$ to lighten the notation.

Proposition 80. *$ran(A - \lambda)$ is a closed subspace of H.*

Corollary 12. *$ran(A - \lambda)$ is a Hilbert space.*

Proposition 81.

$$H = ran(A - \lambda) \oplus N(A - \lambda)$$

And any $\psi \in H$ can thus be written $\psi = \psi_\parallel + \psi_\perp$ for $\psi_\parallel \in ran(A - \lambda)$ and $\psi_\perp \in N(A - \lambda)$.

Proof. Since $\operatorname{ran}(A-\lambda)$ is closed, and $\operatorname{ran}(A-\lambda)^\perp = N(A-\lambda)$, we are done. □

Lemma 3. *For any $\phi \in D_A$, $\phi_\parallel \in \Delta$.*

Proof. Since $\phi \in D_A$ and $\phi_\perp \in N(A - \lambda) \subseteq D_A$, $\phi_\parallel = \phi - \phi_\perp \in D_A$ by closure. □

Lemma 4. *$ran(A - \lambda)|_\Delta = ran(A - \lambda)$*

Proof. Inclusion from left to right holds trivially. Now, let $\psi \in \operatorname{ran}(A - \lambda)$. Then $\psi = A - \lambda\phi$ for $\phi \in D_A$. Since $\phi \in H, \phi = \phi_\parallel + \phi_\perp$. Since $\phi_\parallel \in D_A$ we can use linearity of $(A - \lambda)$ to establish that

$$(A - \lambda)\phi = (A - \lambda)(\phi_\parallel + \phi_\perp) = (A - \lambda)\phi_\parallel$$

But since $\phi_\parallel \in \Delta$ we have that $\psi \in \operatorname{ran}(A - \lambda)|_\Delta$. Therefore inclusion holds both ways and the sets are equal. □

Proposition 82. *The operator $(A - \lambda)|_\Delta :\longrightarrow ran(A - \lambda)$ is 1-1 and onto.*

Proof. Suppose that

$$(A - \lambda)\phi_1 = (A - \lambda)\phi_2$$

for $\phi_1, \phi_2 \in \Delta$. Then

$$A(\phi_1 - \phi_2) = \lambda(\phi_1 - \phi_2)$$

So $\phi_1 - \phi_2 \in N(A - \lambda)$. But ϕ_1 and ϕ_2 are in $\operatorname{ran}(A - \lambda)$, so $\phi_1 - \phi_2 = 0$. So $\phi_1 = \phi_2$.

Onto-ness follows from Lemma 4. □

Corollary 13. *The linear operator* $[(A-\lambda)|_\Delta]^{-1} : ran\,(A-\lambda) \longrightarrow \Delta$ *exists.*

Proposition 83. Δ *is dense in* $ran\,(A-\lambda)$.

Proof. Let $\psi \in \text{ran}\,(A-\lambda)$. Since D_A is dense in H, we have $\exists \psi_n \in D_A$ such that $\lim \psi_n = \psi$. Since $(\psi_n)_\parallel \in \Delta$, and for each n it holds that

$$\|(\psi_n)_\parallel - \psi\| \leq \|\psi_n - \psi\|$$

it follows that $\lim(\psi_n)_\parallel = \psi$. □

Proposition 84. *The operator* $(A-\lambda)|_\Delta : \Delta \longrightarrow ran\,(A-\lambda)$ *is self-adjoint.*

Proof. We use Δ^* to denote the domain of the adjoint.

$$\Delta^* = \left\{ \psi \in \text{ran}\,(A-\lambda) \middle| \exists \eta \in \text{ran}\,(A-\lambda) : \forall \phi \in \Delta, \langle \psi, (A-\lambda)\phi \rangle = \langle \eta, \phi \rangle \right\}$$

Since regular $A - \lambda : D_A \longrightarrow H$ is self-adjoint, we have that $(A-\lambda)|_\Delta$ is symmetric. Therefore $\Delta \subseteq \Delta^*$. It remains to show then that $\Delta^* \subseteq \Delta$. Observe that $\forall \alpha \in D_A, \alpha = \alpha_\parallel + \alpha_\perp$, so that:

$$(A-\lambda)\alpha = (A-\lambda)\alpha_\parallel$$

and therefore for any $\xi \in H$,

$$\langle \xi, (A-\lambda)\alpha \rangle = \langle \xi, (A-\lambda)\alpha_\parallel \rangle$$

Now let $\psi \in \Delta^*$. Then $\exists \eta \in \text{ran}\,(A-\lambda) : \forall \phi \in \Delta$,

$$\langle \psi, (A-\lambda)\phi \rangle = \langle \eta, \phi \rangle$$

But now let $\alpha \in D_A$. Then

$$\langle \psi, (A-\lambda)\alpha \rangle = \langle \psi, (A-\lambda)\alpha_\parallel \rangle = \langle \eta, \alpha_\parallel \rangle$$

But since $\eta \in \text{ran}\,(A-\lambda)$ and $\alpha_\perp \in N(A-\lambda)$, we have

$$= \langle \eta, \alpha_\parallel \rangle + 0 = \langle \eta, \alpha_\parallel \rangle + \langle \eta, \alpha_\perp \rangle = \langle \eta, \alpha_\parallel + \alpha_\perp \rangle = \langle \eta, \alpha \rangle$$

so that reading just the left and right end of this equation chain we obtain

$$\langle \psi, (A-\lambda)\alpha \rangle = \langle \eta, \alpha \rangle$$

But this η must work $\forall \alpha \in D_A$. Therefore $\psi \in D_{A^*}$. But A is self-adjoint, so $\psi \in D_A$. Therefore if $\psi \in \Delta^*$ it must be in $\text{ran}\,(A-\lambda)$ and it must also be in D_A. So $\psi \in \Delta$. □

Corollary 14. $(A-\lambda)|_\Delta$ *and its inverse are closed.*

Corollary 15. $(A-\lambda)|_\Delta^{-1}$ *is bounded.*

Proof. Closed graph theorem. Can be found in [6]. □

Lemma 5. *For any $\epsilon > 0$, $(A - \lambda - i\epsilon)(\Delta) = ran\,(A - \lambda)$.*

Proof. Let $\psi \in LHS$. Then $\psi = (A - \lambda)\phi - i\epsilon\phi$. Since $\phi \in \Delta$, we have $\phi \in \text{ran}\,(A - \lambda)$, so $i\epsilon\phi$ stays in the range, So $\psi \in RHS$.

Now, let $\psi \in RHS$. Then $\psi = (A - \lambda - i\epsilon)\phi$ for some $\phi \in D_A$. But $\phi = \phi_{\|} + \phi_{\perp}$, and so

$$\psi = (A - \lambda - i\epsilon)\phi_{\|} + i\epsilon\phi_{\perp}$$

But $(A-\lambda-i\epsilon)\phi_{\|} \in \text{ran}\,(A-\lambda)$, and $i\epsilon\phi_{\perp} \in N(A-\lambda)$. Since the decomposition is unique, $\phi_{\perp} = 0$, and $\psi = (A - \lambda - i\epsilon)\phi_{\|}$. Since $\phi_{\|} \in \Delta$, $\psi \in LHS$. □

Proposition 85. *For any $\psi \in ran\,(A - \lambda)$,*

$$\lim_{\epsilon \to 0} (A - \lambda - i\epsilon)^{-1}\psi = (A - \lambda|_{\Delta})^{-1}\psi$$

Proof. Lemma 2 □

Proposition 86. *Let $\lambda \in \sigma(A)$ for A with only pure point spectrum. (clean up this notation. Choose a and b such that λ is the only eigenvalue in (a, b). Let $\psi \in ran\,(A - \lambda)$. Then*

$$\lim_{\epsilon \to 0} \int_a^b Im\,\langle \psi, (A - t - i\epsilon)^{-1}\psi \rangle\, dt = 0$$

Proof. See Propositions 79 and 85 □

Proposition 87. *Let $\lambda \in \sigma(A)$ for self-adjoint A with only pure point spectrum. Let $\lambda \in \sigma(A)$. Let $\psi \in N(A - \lambda)$. Then*

$$\mu_{\psi} = \|\psi\|^2 \chi_{\{\lambda\}}$$

Proof. Follows analagously to the proof for the PVM of the identity. □

Proposition 88. *Let $\lambda \in \sigma(A)$ for self-adjoint A with only pure point spectrum. Let $\psi \in H$. Suppose a and b are chosen such that λ is the only eigenvalue in (a, b). Then*

$$\mu_{\psi}(a, b) = \|P_{\lambda}\psi\|^2$$

Proof. Let $\psi \in H$. Then $\psi = \psi_{\|} + \psi_{\perp}$, for $\psi_{\|} \in \text{ran}\,(A-\lambda)$ and $\psi_{\perp} \in N(A-\lambda)$.

Now,

$$\begin{aligned}\mu_\psi(a,b) &= \lim_{\epsilon\to 0}\int_a^b \langle\psi,(A-t-i\epsilon)^{-1}\psi\rangle dt \\ &= \lim_{\epsilon\to 0}\int_a^b \langle\psi_\parallel+\psi_\perp,(A-t-i\epsilon)^{-1}(\psi_\parallel+\psi_\perp)\rangle dt \\ &= \lim_{\epsilon\to 0}\int_a^b \langle\psi_\parallel,(A-t-i\epsilon)^{-1}\psi_\parallel\rangle + \langle\psi_\perp,(A-t-i\epsilon)^{-1}\psi_\perp\rangle \\ &= \lim_{\epsilon\to 0}\int_a^b \langle\psi_\parallel,(A-t-i\epsilon)^{-1}\psi_\parallel\rangle + \lim_{\epsilon\to 0}\int_a^b \langle\psi_\perp,(A-t-i\epsilon)^{-1}\psi_\perp\rangle \\ &= \|\psi_\perp\|^2 = \|P_\lambda\psi\|^2 \quad (7.18)\end{aligned}$$

□

Proposition 89. *Suppose $A : D_A \longrightarrow H$ is a self-adjoint operator with only pure point spectrum and eigenvalues which can be written as an increasing sequence $\lambda_1, \lambda_2, ...$ which is not bounded above. Then*

$$P_A(E) = \sum_{\lambda\in E} P_\lambda$$

Proof. Let $\psi \in H$. Break up the real line into intervals $I_1 = (-\infty, \lambda_1]$, $I_k = (\lambda_{k-1}, \lambda_k]$ for $k = 2, 3, ...$ First, we will show that

$$\bigcup_{k=1}^{\infty} I_k = \mathbb{R},$$

Inclusion from left to right holds trivally. To show inclusion in the other direction, let $x \in \mathbb{R}$. Then since the λ_k sequence is increasing and not bounded above, we have that $\exists N : \lambda_N > x$ whenever $n > N$. Let M be the smallest integer that satisfies this condition. Then $x \in (\lambda_{M-1}, \lambda_M]$. Therefore $x \in LHS$. To calculate an arbitrary P(E) then, we may intersect E with the whole real line:

$$\begin{aligned}\langle\psi, P_A(E)\phi\rangle = \mu_{\psi,\phi}(E) = \mu_{\psi,\phi}(E\cap\mathbb{R}) = \mu_{\psi,\phi}(E\cap(\bigcup_{k=1}^{\infty} I_k)) = \mu_{\psi,\phi}(\bigcup_{k=1}^{\infty}(E\cap I_k)) \\ = \sum_{k=1}^{\infty}\mu_{\psi,\phi}(E\cap I_k) = \sum_{\lambda\in E}\langle\psi, P_\lambda\phi\rangle = \langle\psi, \Big(\sum_{\lambda\in E} P_\lambda\Big)\phi\rangle \quad (7.19)\end{aligned}$$

□

Proposition 90. *Suppose a quantum system $\mathcal{H}$ is in the state ρ_ψ. Suppose an observable $A : D_A \longrightarrow \mathcal{H}$ has eigenvalues satisfying the conditions in Proposition 89, with eigenvectors ψ_n forming an ON basis, so we can write $\psi = \sum_{n=1}^{\infty} c_n\psi_n$. Then the probability of measuring the observable in the set $E \subseteq \mathbb{R}$ is given by*

$$Prob\,(E) = \sum_{\lambda_n\in E} |c_n|^2 \quad (7.20)$$

Proof.

$$\begin{aligned}\text{Prob}\,(E) = \text{Trace}\,(P_A(E)\rho_\psi) = \langle\psi, P_A(E)\psi\rangle = \langle\psi, \Big(\sum_{\lambda_n\in E} P_{\lambda_n}\Big)\psi\rangle \\ = \langle\psi, \sum_{\lambda_n\in E} c_n\psi_n\rangle = \sum_{\lambda_n\in E} c_n\langle\psi,\psi_n\rangle = \sum_{\lambda_n\in E}|c_n|^2 \quad (7.21)\end{aligned}$$

□

Proposition 91. *If a quantum system with Hamiltonian H is in a pure state ρ_ψ initially, then after time t it will be in state $\rho_{U_t\psi}$.*

Proof. According to Axiom 5, the state after time t will be

$$\rho_t = U_t\rho_\psi U_t^*$$

Now, let $\phi \in H$. Then

$$\rho_t\phi = U_t\rho_\psi U_t^*\phi = U_t\langle\psi, U_t^*\phi\rangle\psi = \langle\psi, U_t^*\phi\rangle U_t\psi = \langle U_t\psi,\phi\rangle U_t\psi = \rho_{U_t\psi}\phi$$

□

8. OBSERVABLES AND TIME EVOLUTION

How, then, does one construct the operators associated with observables quantities in nature? A naïve strategy is this: consider first our motive to define a position operator associated with my coordinate x. If I have a wave function, the most obvious guess of "multiplying" the position operator with a wave function is to multiply the coordinate value by the wave function. That is, $Q\psi(x)$ should be $x\psi(x)$. In order to use many of my theorems, however, I require that Q should be self-adjoint. As we may recall, self-adjointness is a consequence of the domain of definition of an operator. How do we define the domain of definition? In the case of a particle on the real line, we will have the Hilbert space of the particle (at first ignoring spin) is $H = L^2(\mathbb{R})$. Suppose first we could define Q on the whole Hilbert space. Observe that the function f defined by $f(x) = 1/x$ for $x > 1$ and $f(x) = 0$ otherwise is in $L^2(\mathbb{R})$. Now, $Qf(x) = 1$ for $x > 1$, and so $\|Qf\|^2 = \int_1^\infty 1dx = \infty$, so Qf doesn't even land in H. This position operator with domain as the entire Hilbert space thus does not have range in H. The appearance of an unwanted infinity for $\|Qf\|$ is a harbinger of the fact that the correctly defined position operator will be unbounded. For now, we toss out this Q as our candidate.

We instead approach the construction of observables from a different perspective. It is not necessarily the most intuitive approach, but it is very convenient in that it guarantees that the domain is self-adjoint.

his approach is to define our observables as the generators of a special one parameter abelian group of unitary operators on the Hilbert space of the system.

The following theorem is a beautiful way to obtain the observables with their domains for free. We do not provide its proof, but it can be found in (reference).

Theorem 8. *(Stone's theorem) Let $\{U_t | t \in \mathbb{R}\}$ be a strongly continuous one-parameter unitary group. Then there exists a self-adjoint operator $A : D_A \longrightarrow H$ such that $\forall t \in \mathbb{R}$,*

$$U_t = e^{itA}$$

and the operator $A : D_A \longrightarrow H$ is given by

$$D_A = \{\psi \in H | \lim_{\epsilon \to 0} -\frac{i}{\epsilon}(U_\epsilon(\psi) - \psi) \text{ exists}\}$$

and

$$A\psi = -i \lim_{\epsilon \to 0} \frac{U(\epsilon)\psi - \psi}{\epsilon}$$

Proof. Can be found in **??**. □

Corollary 16. *There is a one to one correspondence between unitary groups of the type describe in Theorem 8 and self-adjoint operators on a Hilbert space.*

8.1 Momentum

From classical mechanics, we know that momentum is the generator of translations. In quantum mechanics for particles on the real line, we $define$ the momentum operator as the generator of the group of translations.

Definition 87. *The group of translations $\{U_t^P\}_{t\in\mathbb{R}}$ is defined by*

$$U_t^P\psi(x) = \psi(x-t)$$

Definition 88. *The momentum operator $P : D_P \longrightarrow \mathcal{H}$ is defined as the generator of U_t^P, in units such that $\hbar = 1$.*

Proposition 92. *On the Schwartz space $\mathcal{S}$, it is true that*

$$P\psi = -i\psi'$$

8.2 Position

The position operator is defined similarly as the generator of a unitary group of transformations on the Hilbert space. It is perhaps not as intuitive as for momentum, but it's what we want:

Definition 89. *The group of phases $\{U_t^Q\}_{t\in\mathbb{R}}$ is defined by*

$$U_t^Q\psi(x) = e^{itx}\psi(x)$$

Remark 18. *Each U_t effectively winds up a wave function ψ, with larger values of the parameter t corresponding to tighter and tighter coils.*

Definition 90. *The position operator $Q : D_Q \longrightarrow \mathcal{H}$ is defined as the generator of U_t^Q.*

Proposition 93. *On the Schwartz space $\mathcal{S}$, it is true that*

$$Q\psi(x) = x\psi(x)$$

Proof. Can be found in [9]. □

8.3 The Fourier Transform

The Fourier transform shows up in a perhaps unprecedented way in quantum mechanics. Often, it is used to talk about the "frequency distribution" of a time series signal. As nature would have it, position and momentum distributions in quantum mechanics follow the same rule, which we shall state precisely in the next section.

Definition 91. *The Fourier operator $\mathcal{F} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$ is defined for $\psi \in mathcalS(\mathbb{R})$ by*

$$F(\psi)(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) e^{-ixy} dy$$

It can be shown that this is a unitary operator and then can be extended to the whole Hilbert space. The proof of this can be found in [9]

8.4 The relationship of the position and momentum operators

The position and momentum operators are related by a Fourier transform.

Proposition 94. *Let $P : D_P \longrightarrow H$ denote the momentum operator, $Q : D_Q \longrightarrow H$ denote the position operator, and $U : \mathcal{H} \longrightarrow \mathcal{H}$ denote the Fourier tranform operator. Then*

$$P = UQU^{-1} \tag{8.1}$$

Proof. Can be found in [9]. □

8.5 Time Evolution

There is one self-adjoint operator in particular which we do not define as the generator of an already defined group, but is defined a priori, and determines an extremely important one parameter group. This operator is the Hamiltonian, and the one parameter group generated by the Hamiltonian is the group of time evolution operators.

Definition 92. *The group of time evolution operators is given by*

$$U_t = e^{\frac{-iHt}{\hbar}}$$

where $e^{\frac{-iHt}{\hbar}}$ is defined according to the spectral theorem for unbounded operators.

Proposition 95. *If a quantum system with Hamiltonian H is in a pure state ρ_ψ initially, then after time t it will be in state $\rho_{U_t\psi}$.*

Proof. According to Axiom 5, the state after time t will be

$$\rho_t = U_t \rho_\psi U_t^*$$

Now, let $\phi \in H$. Then

$$\rho_t \phi = U_t \rho_\psi U_t^* \phi = U_t \langle \psi, U_t^* \phi \rangle \psi = \langle \psi, U_t^* \phi \rangle U_t \psi = \langle U_t \psi, \phi \rangle U_t \psi = \rho_{U_t \psi} \phi$$

□

9. THE HARMONIC OSCILLATOR

Now that we have developed a complete formalism for dealing with quantum mechanical systems, we may apply it to a simple quantum mechanical system that has widespread applications. One may recall from classical mechanics that the Hamiltonian of the harmonic oscillator is

$$H(p,q) = \frac{p^2}{2m} + \frac{1}{2}m\omega q^2$$

One would like to translate a classical H into a quantum mechanical system by the following transformation:

$$q \longrightarrow Q, \qquad p \longrightarrow P$$

in order to write down the operator equation:

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega Q^2 \tag{9.1}$$

This is, in spirit, what we do. However, there are two substantial technical considerations we must take care of. Equation 9.1 is in fact incorrect.

In the first place, we must worry about domain. The domains of the P and Q operators are different, and the domains of P^2 and Q^2 differ even more, so we could take care to make sure define the domain of H to be the intersection of the domains of the squares, but we have no reason to believe this would make H self-adjoint. To take care of this, we use the Schwartz space. First we denote the restrictions of the position and momentum operators to the Schwarz domain by $\widetilde{Q}$ and $\widetilde{P}$. Now, since we know the Schwarz space is invariant under Q and P, may define the operator

$$\widetilde{H} := \frac{\widetilde{P}^2}{2m} + \frac{1}{2}m\omega\widetilde{Q}^2 \tag{9.2}$$

Now, having restricted ourselves to the Schwartz space, we make note of the issue with the well-definition of this transformation from a classical Hamiltonian to a quantum Hamiltonian. First, consider any two classical Hamiltonians H(p,q) and H'(p,q). Consider now the Hamiltonian

$$H''(p,q) = H'(p,q) + \frac{(pq - qp)}{i\hbar}(H(p,q) - H'(p,q))$$

Now, define as previously the quantum analogues, this time adding a hat above each one to distinguish them from their classical counterparts. We would

have in the classical case that $H = H''$, so that if such a mapping were well-defined, $\hat{H} = \hat{H''}$. But $\hat{H''} = \hat{H'}$, so $\hat{H} = \hat{H'}$. Therefore all quantum hamiltonians generated in this way are equal. As to why someone would so pathologically add this, we are not concerned. The important takeaway is that it presents inconsistencies, and we therefore cannot trust this technique religiously.

The misstep in logic here was that H and H'' are not mapped to the same quantum Hamiltonian, despite being equivalent as functions. Therefore, we do not wish here to motivate the transition from quantum to classical, we only start with the operator Equation 9.2 and try to build a self-adjoint Hamiltonian from it.

It can be shown that $\widetilde{H}$ is symmetric, and by solving the eigenvalue problem one obtains a sequence of eigenvalues for H,

$$E_n = (n + \frac{1}{2})\hbar\omega, \qquad n = 0, 1, 2$$

with eigenvectors

$$\psi_n = H_n\Big(\sqrt{\frac{m\omega}{2\hbar}}x\Big)\exp\Big(\sqrt{\frac{m\omega}{2\hbar}}x\Big)$$

when $H_n(\xi)$ is the nth Hermite polynomial. It can also be shown the that ψ_n's above form an orthonormal basis. This proof can be found in [1]

Theorem 9. *If the eigenvectors of a symmetric operator $A : D_A \longrightarrow H$ form an orthonormal basis, then A is essentially self-adjoint.*

Corollary 17. *$\widetilde{H}$ as defined above is essentially self-adjoint.*

We then define the quantum harmonic oscillator Hamiltonian H as the unique self-adjoint extension of $\widetilde{H}$.

9.1 Measurement Statistics of a Pure State

Suppose we have a particle in a harmonic potential in the pure state ρ_ψ.

9.1.1 Position and Momentum

We restate the following two results for completeness. Although we have already provided the proof of these statements, we will restate the general result for a specific case. If a particle in the quantum harmonic oscillator is in the pure state ρ_ψ, then the probability of measuring the particle's position as being in the set E is

$$Prob(E) = \int_E |\psi(x)|^2 dx$$

and the probability of obtaining a momentum value in E is:

$$Prob(E) = \int_E |\hat{\psi}(p)|^2 dp$$

where $\hat{\psi}$ is the Fourier transform of f.

9.1.2 Energy

Let E be a measurable subset of the real line. According to Axiom 4, the probability of measuring the energy as being in the set E is

$$Prob(E) = Tr(\rho_\psi P_H(E)) = \langle \psi, P_H(E)\psi \rangle = \sum_{\lambda_i \in E} |c_i|^2$$

Thus, all of the ordinary probability rules that used to seem separate are all contained in the statement of Axiom 4.

9.2 Time Evolution

From Propostion 91, we know that if a system starts in state ρ_ψ, then the pure state after time t is given by $\rho_{U_t\psi}$. So, let $\psi = \sum c_n \psi_n$ where ψ_n is the basis of eigenfunctions for the harmonic oscillator. Then

$$U_t\psi = e^{\frac{-iHt}{\hbar}}\psi = e^{\frac{-iHt}{\hbar}}\sum_n c_n\psi_n = \sum_n c_n e^{\frac{-iHt}{\hbar}}\psi_n = \sum_n c_n \int e^{\frac{-i\lambda t}{\hbar}}\, dP\, \psi_n = \sum_n c_n e^{\frac{-iE_n t}{\hbar}}\psi_n \tag{9.3}$$

Thus, the ordinary time evolution that may be found in [1] for the Harmonic oscillator can be viewed as a consequence of Axiom 5.

10. CONCLUSION

I would like to reiterate the statement that much of the understanding developed here is not spectacularly useful for practical calculations of most quantum mechanical systems. We have seen how long it takes to prove the equivalence between our theory and the usual approach.

However, for a student interested in truth, this development of the axioms is paramount. We can finally say that there *are* unambiguous laws of quantum mechanics. Furthermore, if we want to go back to the usual approach, we understand what it really is that we are saying about a system, and we know where we're making shortcuts and where we're stating the truth.

In the future, I hope to write a paper to explore the sixth axiom of projective dynamics, especially in the context of quantum computers.

There is a certain beauty in stating the laws of the universe so generally and carefully, which cannot be appreciated when making ad-hoc modifications to theorems depending on such things as the type of spectrum of an operator. I hope the reader will find that this paper demonstrated this beauty effectively, and will aid them in their desire to understand the inner-workings of the universe.

BIBLIOGRAPHY

[1] David J. Griffiths. Griffiths, D. *Introduction to Quantum Mechanics*. 2nd ed., Cambridge University Press, 2017.

[2] Dirac, P.A.M. *The Principles of Quantum Mechanics*. 4th ed., Oxford Science Publication. 1958. *The Principles of Quantum Mechanics*

[3] Sakurai, J. Napoloitano, J. *Modern Quantum Mechanics*. 2nd ed., Cambridge University Press, 2017.

[4] Shankar, R. *Principles of Quantum Mechanics*. 2nd ed., Springer, 2014.

[5] Schuller, Frederic. Youtube. "Axioms of Quantum Mechanics - Lec01 - Frederic Schuller". 12 March 2016.

[6] Akhiezer, N. Glazman, I. *Theory of Linear Operators in Hilbert Space*. Dover Publications Inc. 1980.

[7] Royden, H. *Real Analysis*. 3rd ed., Macmillan, 1988. *Real Analysis*

[8] Rudin W. *Real and Complex Analysis*. 3rd ed., McGraw-Hill, 1987.

[9] Reed, M. Simon, B. *Methods of Modern Mathematical Physics: Functional Analysis*. Academic Press, 1972.