Hilbert Spaces

Felix Flicker

Table of contents

Introduction

Welcome to the Hilbert Spaces notes!

Hilbert Spaces were introduced to formalise the mathematics behind quantum mechanics. Bristol's most famous alumnus, Paul Dirac, used Hilbert spaces to show the equivalence between the wave mechanics of Schroedinger to the matrix mechanics of Heisenberg, thus creating the unified theory of quantum mechanics. We will look at the mathematics, philosophy, and history of these ideas.

These notes are a short summary of the key axioms and derivations for the Hilbert Spaces section of the Advanced Mathematics for Physicists course. They are very much intended to accompany, not replace, the lectures!

Course Overview

The chapters leading up to Chapeter Chapter 9 introduce the mathematics of Hilbert spaces and a more mathematical look at the underpinning of quantum mechanics.

In Chapter Chapter 10 we take a detailed look at the respective positions of Dirac and von Neumann about the relevance of Hilbert spaces to quantum mechanics.

We then use this newfound understanding to look at orthogonal polynomials in Chapters Chapter 11 to Chapter 17. In this way you will see the deep connections between waves, discrete states, and their description in Hilbert spaces.

References

- John von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press 1955 (translated and updated from the original German edition of 1932)
- Paul Adrien Maurice Dirac, *Principles of Quantum Mechanics*, Oxford University Press 1930 (but get the revised 4th edition of 1967: he invents Dirac notation in the 3rd edition)
- Richard Courant and David Hilbert, *Methods of Mathematical Physics I,* Gottingen 1924 (get the 2nd edition, 1953)
- Gábor Szegő, *Orthogonal Polynomials,* AMS 1973
- [Stanford Encyclopedia of Philosophy] Quantum Theory: von Neumann vs Dirac

Notation

This is a list of all the notation we will be using in this part of the course. It is all standard, and part of the purpose of this course is to familiarise you with symbols you will see in other courses but which might not be explained.

- $\triangleq$: equal by definition (technically standard notation, but you may be more familiar with $\equiv$ for this. But $\equiv$ also means 'equivalent to').
- $|v\rangle$: a vector.
- $\{a, b, c\}$: the set containing a,b, c (ordering doesn't matter)
- (a, b, c): the sequence a, b, c (ordering matters)
- (a, b): an ordered pair (sequence of two elements)
- $\forall$: for all
- $\exists$: there exists
- $|$: such that
- $a \in S$: a is a member of the set S
- $S \subset T$: S is a strict subset of the set T
- $S \subseteq T$: S is a subset of the set T (*i.e.* it may be the same set)
- $A \times B$: cartesian product
- $f : S \times T \to U$: *binary function* f takes in an element of S and an element of T and returns an element of U
- $f : S \times S \to S$: *binary operation* (binary function where all sets are the same set)
- $a \implies b$: a implies b
- iff: if and only if
- $a \iff b$: a iff b; equivalently, a implies b and b implies a
- $\mathfrak{Re}(z)$: real part of z
- $\mathfrak{Im}(z)$: imaginary part of z

- $\aleph_0$ (aleph-null): the cardinality of the integers, *a.k.a.* countable infinity
- $\mathfrak{c}$: the cardinality of the continuum

1 Fields

1.1 Cartesian products, binary functions, binary operations

A *Cartesian product* $A \times B$ is defined to be

$$A \times B = \{(a, b) \,|a \in A, b \in B\}$$

i.e. it is the set of all ordered pairs of elements. The original example is the real 2D plane $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. To see this, note that each point in the plane can be assigned Cartesian co-ordinates (x, y). This is an ordered pair, since $(x, y) \neq (y, x)$ (order matters).

A binary function is any function that takes in two arguments. We denote this

$$f : A \times B \to C$$

meaning f takes in elements of A and B and returns elements in C. For example, $\alpha \cdot |v\rangle = |v'\rangle$, where α is a number, $|v\rangle$ and $|v'\rangle$ are vectors, and $\cdot$ is multiplication. We can denote this

$$\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V}.$$

A binary operation is a binary function in which all three sets are the same:

$$f : A \times A \to A.$$

For example, multiplying together two real numbers gives another real number:

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}.$$

1.2 Definition of a Field

A field $\mathcal{F}$ is a non-empty set F along with the binary operations

$$+ : F \times F \to F \text{ (\textit{field addition})}$$

and

$$\cdot : F \times F \to F \text{ (\textit{field multiplication})}$$

obeying the following axioms.

$\forall \alpha, \beta, \gamma \in F :$

- [**F1a**]: $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ (*associativity of* $+$)
- [**F1b**]: $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$ (*associativity of* $\cdot$)
- [**F2a**]: $\alpha + \beta = \beta + \alpha$ (*commutativity of* $+$)
- [**F2b**]: $\alpha \cdot \beta = \beta \cdot \alpha$ (*commutativity of* $\cdot$)
- [**F3a**]: $\exists 0 \in F | \alpha + 0 = \alpha$ (*additive identity*)
- [**F3b**]: $\exists 1 \in F | \alpha \cdot 1 = \alpha$ (*multiplicative identity*)
- [**F3c**]: $0 \neq 1$ (*non-triviality* – **my name!**)
- [**F4a**]: $\exists (-\alpha) \in F | \alpha + (-\alpha) = 0$ (*additive inverse*)
- [**F4b**]: $\forall (\alpha \neq 0) \exists \alpha^{-1} \in F | \alpha \cdot \alpha^{-1} = 1$ (*multiplicative inverse*)
- [**F5**]: $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$ (*distributivity of* $\cdot$ *over* $+$).

Alternatively, we can say that a field is an *ordered pair* $\mathcal{F} = (F, \{+, \cdot\})$ obeying the axioms above.

1.3 Examples of Fields

Examples include:

- $\mathbb{R}$ (the reals)
- $\mathbb{Q}$ (the rationals)
- $\mathbb{C}$ (complex numbers)

2 Vector Spaces

Sometimes called *linear spaces*: vector spaces are linear by definition.

2.1 Definition of a Vector Space

A vector space $\mathcal{V}$ over a field $\mathcal{F}$ is a non-empty set V together with a binary operation

$$\oplus : V \times V \to V \ (\textit{vector addition})$$

and a binary function

$$\otimes : F \times V \to V \ (\textit{scalar multiplication})$$

[**⊕, ⊗ are my notation!**] obeying the following axioms.

$\forall |a\rangle, |b\rangle, |c\rangle \in V; \ \alpha, \beta \in F$:

- **[V1]**: $|a\rangle \oplus |b\rangle \in V$ (*closure under* $\oplus$)
- **[V2]**: $|a\rangle \oplus |b\rangle = |b\rangle \oplus |a\rangle$ (*commutativity of* $\oplus$)
- **[V3]**: $(|a\rangle \oplus |b\rangle) \oplus |c\rangle = |a\rangle \oplus (|b\rangle \oplus |c\rangle)$ (*associativity of* $\oplus$)
- **[V4]**: $\exists |0\rangle \in V | |a\rangle \oplus |0\rangle = |a\rangle$ (*identity element under* $\oplus$)
- **[V5]**: $\exists | -a\rangle \in V | |a\rangle \oplus | -a\rangle = |0\rangle$ (*inverse elements under* $\oplus$)
- **[V6]**: $\alpha \otimes |a\rangle \in V$ (*closure under* $\otimes$)
- **[V7]**: $\alpha \otimes (\beta \otimes |a\rangle) = (\alpha \cdot \beta) \otimes |a\rangle$ (*compatibility of* $\otimes$ *with* $\cdot$)
- **[V8]**: $\exists 1 \in F | 1 \otimes |a\rangle = |a\rangle$ where 1 is the multiplicative identity in $\mathcal{F}$ (*identity element under* $\otimes$)
- **[V9]**: $\alpha \otimes (|a\rangle \oplus |b\rangle) = \alpha \otimes |a\rangle \oplus \alpha \otimes |b\rangle$ (*distributivity of* $\otimes$ *w.r.t.* $\oplus$)
- **[V10]**: $(\alpha + \beta) \otimes |a\rangle = \alpha \otimes |a\rangle \oplus \beta \otimes |a\rangle$ (*distributivity of* $\otimes$ *w.r.t.* $+$).

Equivalently, a vector space is an ordered pair $\mathcal{V} = (V, \{\otimes, \oplus\})$ obeying these axioms.

NB the zero vector $|0\rangle$ should not be confused with the 0 element of the field.

V7 establishes that no ambiguity is introduced if we denote scalar multiplication $\otimes$ with the same symbol $\cdot$ used for field multiplication. Similarly, **V10** establishes that $\oplus$ can be written $+$ without introducing any confusion. Strictly the two binary operations play different roles, but I have never seen them distinguished symbolically in the literature. Hence, from hereon I will use $\cdot$ for both field multiplication and scalar multiplication (or will omit the symbol all together, as is customary), and $+$ for both field addition and vector addition.

2.2 Examples of Vector Spaces

- $\{|0\rangle\}$, the set containing only the zero vector
- $\mathcal{F}$, the field over which $\mathcal{V}$ is defined (e.g. $(\mathbb{R}, \{+, \cdot\})$)
- $\mathcal{F}^N$, the $N-$dimensional co-ordinate space. This is perhaps the most familiar example. *E.g.*

$$|x\rangle = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} ; \quad |x\rangle + |y\rangle = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ \vdots \\ x_N + y_N \end{pmatrix} \ (etc.)$$

with the absolutely familiar case being $\mathbb{R}^N$.

3 Inner Product Spaces

3.1 Notations for Inner Products

Inner products are variously denoted

$$\langle\cdot|\cdot\rangle; \quad \langle\cdot,\cdot\rangle; \quad (\cdot,\cdot)\,; \quad (\cdot|\cdot)\,.$$

I will principally use the first of these in order to connect to Dirac notation familiar from quantum mechanics. I will sometimes use the second.

3.2 Definition of an Inner Product Space

An Inner Product Space $\mathcal{I}$ is a vector space $\mathcal{V}$ over the field $\mathcal{F} = \mathbb{R}$ or $\mathbb{C}$ together with a binary function

$$\langle\cdot|\cdot\rangle : V \times V \to F \text{ (the \textit{inner product})}$$

obeying the following axioms.

$\forall |a\rangle, |b\rangle, |c\rangle \in V;\ \alpha, \beta \in F$:

- **[I1]**: $\langle a|b\rangle = (\langle b|a\rangle)^*$ (*conjugate symmetry*)
- **[I2]**: $\langle c|\,(\alpha|a\rangle + \beta|b\rangle) = \alpha\langle c|a\rangle + \beta\langle c|b\rangle$ (*linearity in the second argument*)
- **[I3]**: $\langle a|a\rangle \geq 0$, & $(\langle a|a\rangle = 0) \iff (|a\rangle = |0\rangle)$ (*positive definiteness*).

Equivalently, an Inner product space is an ordered pair $\mathcal{I} = (\mathcal{V}, \langle\cdot|\cdot\rangle)$ obeying these axioms.

NB in the mathematics literature, linearity in the first argument is often used instead. The results are equivalent, but a consistent choice must be made.

3.3 Examples of Inner Product Spaces

- $\mathbb{C}$, with $\langle x|y\rangle = x^* y$.
- complex functions $f(x)$, with $\langle f|g\rangle = \int f^*(x)\, g(x)\, \mathrm{d}x$.
- $\mathbb{C}^N$ with the Euclidean inner product:

$$|x\rangle = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}, \quad \langle x|y\rangle = \sum_{i=1}^{N} x_i^* y_i.$$

3.4 Properties of Inner Product Spaces

3.4.1 Lemma [lI1]: $\langle a|0\rangle = \langle 0|a\rangle = 0$

Proof:

$$[\mathbf{V4}]: \quad \exists |0\rangle \in V | |a\rangle + |0\rangle = |a\rangle \forall |a\rangle \in V.$$

Choose $|a\rangle = |0\rangle$. Therefore

$$|0\rangle = |0\rangle + |0\rangle.$$

Create inner products with $\langle b|$:

$$\langle b|0\rangle = \langle b| \left(|0\rangle + |0\rangle\right)$$

Now use [**I2**]:

$$\langle b|0\rangle = \langle b|0\rangle + \langle b|0\rangle$$

and subtract $\langle b|0\rangle$ from both sides:

$$0 = \langle b|0\rangle.$$

For the other case, use [**I1**]:

$$0 = (\langle 0|b\rangle)^*$$

and take the complex conjugate of both sides:

$$0 = \langle 0|b\rangle.$$

QED.

3.4.2 Lemma [lI2]: $\langle x|x\rangle \in \mathbb{R}$.

Proof: from [**I1**],

$$\langle x|y\rangle = (\langle y|x\rangle)^*$$

and therefore

$$\langle x|x\rangle = (\langle x|x\rangle)^*$$

QED.

4 Linear Maps

4.1 Definition of a Linear Map

The intuition to hold in mind is that a linear map on any finite dimensional vector space is a matrix.

For any two vector spaces $\mathcal{V}$ and $\mathcal{W}$ defined over the same field $\mathcal{F}$, a function $f : V \to W$ is a *linear map* iff $\forall |a\rangle, |b\rangle \in V; \alpha \in F$:

- **[L1]**: $f\left(|a\rangle + |b\rangle\right) = f\left(|a\rangle\right) + f\left(|b\rangle\right)$ (*additivity*)
- **[L2]**: $f\left(\alpha|a\rangle\right) = \alpha f\left(|a\rangle\right)$ (*homogeneity*).

4.2 Examples of Linear Maps

- An $m \times n$ real matrix A acting on a vector $|a\rangle \in \mathbb{R}^n$ returns a vector $A|a\rangle \in \mathbb{R}^m$ (note the different dimensions of the spaces). The most familiar example is $m = n$.
- A function $f\left(x\right) = \alpha x$ for $\alpha, x \in \mathbb{R}$ is a linear map.
- Derivatives acting on smooth functions, and integrals acting on integrable functions, are both linear maps.

4.3 Key Example: Dual Spaces

A key example of a linear operator is the bra $\langle v|$. In physics we are used to thinking of the bra as the Hermitian conjugate of the vector $|v\rangle$. That is:

$$|v\rangle = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \implies \langle v| = \left(x_1^*, x_2^*, x_3^*, \ldots x_N^*\right).$$

This leads us to think of bras and kets on somewhat equal footing. However, mathematically they are at first sight rather different objects. Specifically, while $|v\rangle$ is a vector:

$$|v\rangle \in V$$

$\langle v|$ is a linear map:

$$\langle v| : V \to \mathbb{C}.$$

I.e. the bra is defined as a map which takes in a vector and returns a complex number (by making an inner product):

$$\langle a|b\rangle \in \mathbb{C}.$$

While a ket, *a.k.a.* vector, $|v\rangle$ lives in a vector space $\mathcal{V}$, the bra $\langle v|$ lives in a 'dual vector space' $\mathcal{V}^*$, abbreviated to 'dual space'. Pleasingly, the dual space itself becomes a vector space if endowed with addition and scalar multiplication, as follows.

4.4 Definition of Dual Spaces

Given a vector space $\mathcal{V}$ defined over a field $\mathcal{F}$, a dual space $\mathcal{V}^*$ is the set V^* of all linear maps $\langle v| : \mathcal{V} \to \mathcal{F}$ combined with the binary operations $+$ (addition) and $\cdot$ (scalar multiplication) obeying the following axioms.

$\forall \langle a|, \langle b| \in V^*;\ |v\rangle \in V;\ \alpha \in F$:

- **[V*1]**: $(\langle a| + \langle b|)\,|v\rangle = \langle a|v\rangle + \langle b|v\rangle$ (*distributivity*)
- **[V*2]**: $(\alpha\langle a|)\,|v\rangle = \alpha\langle a|v\rangle$ (*associativity*).

Hence, dual spaces are vector spaces (given **[V*1]** and **[V*2]**) and we *can* think of bras and kets on equal footing, but this was not obvious *a priori*.

4.5 Self-Adjoint vs Hermitian Operators

For notational convenience in this section, I will denote the inner product $\langle\cdot|\cdot\rangle$ as $\langle\cdot,\cdot\rangle$.

4.5.1 The Hermitian adjoint

Consider an inner product space $\mathcal{I} = (\mathcal{V}, \langle \cdot, \cdot \rangle)$ and a linear map $\hat{A} : \mathcal{I} \to \mathcal{I}$. The *Hermitian adjoint* $\hat{A}^\dagger : \mathcal{I} \to \mathcal{I}$ is defined as follows.

$\forall |a\rangle, |b\rangle \in \mathcal{I}$:

- **[H1]**: $\langle a, \hat{A}b \rangle \triangleq \langle \hat{A}^\dagger a, b \rangle$ (*Hermitian adjoint*).

We can equivalently write this in Dirac notation as:

- **[H1]**: $\langle a| \left(\hat{A}|b\rangle \right) \triangleq \left(\left(\langle b|\hat{A}^\dagger \right) |a\rangle \right)^*$ (*Hermitian adjoint*).

4.5.2 Definition of Self-adjoint Operators

When specifying an operator we must specify the domain on which it acts. In quantum mechanics we specify this as part of the problem.

A *self-adjoint operator* on a complex inner product space $\mathcal{I} = (\mathcal{V}, \langle \cdot | \cdot \rangle)$ is a linear map $A : \mathcal{V} \to \mathcal{V}$ that is its own adjoint with the same domain.

If $\mathcal{V}$ is finite-dimensional, the domain of the operator and its adjoint are always the same, and self-adjoint is equivalent to Hermitian. If the domain of the operator and its adjoint are not the same, the operator is Hermitian but not self-adjoint. See the lectures for examples.

Importantly, the nice properties of self-adjoint operators mentioned below do not necessarily hold for Hermitian operators. For example, self-adjoint operators have real eigenvalues, but Hermitian operators need not (consider the 'eigenstate' $\exp(\kappa x), \kappa \in \mathbb{R}$ of the momentum operator!).

In physics we often use the phrase Hermitian when we really mean self-adjoint. Here's a simple example of the difference.

Consider the momentum operator for a quantum particle in a 1D box. We know that $\hat{p} = -i\hbar\partial/\partial x$. Is it Hermitian? Is it self-adjoint?

For the particle in a box we require wavefunctions to be smooth (infinitely differentiable) and to vanish at the boundaries. Hence the domain of $\hat{p}$, $D(\hat{p})$, is

$$D\left(\hat{p}\right) = \left\{\psi = C^\infty\left[0,1\right] | \psi(0) = \psi(1) = 0\right\}$$

where C^∞ denotes functions that can be differentiated an infinite number of times. The domain of $\hat{p}^\dagger$ is defined to be

$$\phi \in D\left(\hat{p}^\dagger\right) \implies \exists \chi | \langle \psi, \hat{p}\phi \rangle = \langle \chi, \psi \rangle \forall \psi \in D\left(\hat{p}\right).$$

We can check that $\widehat{p}$ is Hermitian:

$$\langle \psi, \widehat{p}\phi \rangle = \int_0^1 \psi^*(x)\,(-i\hbar\phi')\,\mathrm{d}x \tag{4.1}$$

$$= -i\hbar\,[\psi^*\phi]_0^1 + \int_0^1 (-i\hbar\psi^*(x))'\,\phi\mathrm{d}x \tag{4.2}$$

$$= \langle \widehat{p}\psi, \phi \rangle \tag{4.3}$$

where the last line follows because ψ vanishes at the boundaries. But this is true regardless of the form of ϕ. So the domain of the adjoint has no restrictions, other than the requirement of smoothness. Hence,

$$D\left(\widehat{p}^\dagger\right) = \{\psi = C^\infty\,[0,1]\}\,.$$

This is a much bigger set of functions than ψ, as there are no boundary conditions imposed! Hence

$$D\left(\widehat{p}\right) \subset D\left(\widehat{p}^\dagger\right)$$

and so $\widehat{p}$ is not self-adjoint. Hence it does not necessarily have real eigenvalues! Perhaps this is obvious: what functions, obeying the boundary conditions, are eigenstates of $\widehat{p}$?

Now consider a slightly different problem: the particle on a ring. In this case we take the less restrictive boundary conditions

$$D\left(\widehat{p}\right) = \{\psi = C^\infty\,[0,1]\,|\psi(0) = \psi(1)\}$$

where we only require that the wavefunction returns to itself at the ends of the domain (so it is a quantum particle living on a ring). This time you can check that the boundary term only vanishes if $\phi(0) = \phi(1)$. But that's the same condition as is applied to ψ. Hence, this time, $\widehat{p}$ is both Hermitian **and** self adjoint: $D\left(\widehat{p}^\dagger\right) = D\left(\widehat{p}\right)$.

Now we can easily write down normalised eigenfunctions obeying the boundary conditions:

$$\widehat{p}|p_n\rangle = p_n|p_n\rangle$$

with

$$\langle x|p_n\rangle = \exp\left(\pm i\hbar p_n x\right)$$

and

$$\hbar p_n \in 2\pi\mathbb{N}.$$

This makes physical sense if you think about momentum: a particle cannot have a well defined momentum, for all time, if trapped in a box; but it can if it can fly out of one end of the box and into the other.

4.5.3 Properties of self-adjoint operators

- [l**H1**] (*involubility*): $\left(\hat{A}^\dagger\right)^\dagger = \hat{A}$
- [l**H2**] (*compatibility of inverse and Hermitian conjugate*): $\left(\hat{A}^\dagger\right)^{-1} = \left(\hat{A}^{-1}\right)^\dagger$
- [l**H3**] (*conjugate linearity*)
  - [l**H3.1**]: $\left(\hat{A} + \hat{B}\right)^\dagger = \hat{A}^\dagger + \hat{B}^\dagger$
  - [l**H3.2**]: $\left(\alpha\hat{A}\right)^\dagger = \alpha^*\hat{A}^\dagger$
- [l**H4**] (*anti-distributivity*): $\left(\hat{A}\hat{B}\right)^\dagger = \hat{B}^\dagger\hat{A}^\dagger$.

We can also prove two important theorems.

Theorem [tH1] (*Self-adjoint operators have real eigenvalues*):

$$\hat{A} = \hat{A}^\dagger \quad \& \quad \hat{A}|a_n\rangle = a_n|a_n\rangle \implies a_n \in \mathbb{R}.$$

Proof:

consider

$$\langle a_n| \left(\hat{A} - \hat{A}^\dagger\right) |a_n\rangle = 0$$

which follows from the definition of self-adjoint. Hence

$$\langle a_n|\hat{A}|a_n\rangle = \langle a_n|\hat{A}^\dagger|a_n\rangle.$$

Now use [**H1**]:

$$\langle a_n|\hat{A}^\dagger|a_n\rangle = \left(\langle a_n|\hat{A}|a_n\rangle\right)^*$$

to give

$$\langle a_n|\hat{A}|a_n\rangle = \left(\langle a_n|\hat{A}|a_n\rangle\right)^*.$$

The theorem requires that $|a_n\rangle$ is an eigenvector of $\hat{A}$, and so:

$$\begin{aligned} a_n\langle a_n|a_n\rangle &= (a_n\langle a_n|a_n\rangle)^* \\ &= a_n^* \left(\langle a_n|a_n\rangle\right)^*. \end{aligned}$$

From [**I1**]:

$$\langle a_n|a_n\rangle = (\langle a_n|a_n\rangle)^*$$

and so

$$a_n\langle a_n|a_n\rangle = a_n^*\langle a_n|a_n\rangle$$

or

$$a_n = a_n^*.$$

QED

Theorem [tH2] (*eigenvectors of $\hat{A}$ with non-degenerate eigenvalues are orthogonal*):

$$\forall \hat{A}|a_n\rangle = a_n|a_n\rangle, \quad \hat{A} = \hat{A}^\dagger :$$

$$a_n \neq a_m \implies \langle a_n|a_m\rangle = 0.$$

Proof:

Now consider

$$\langle a_n|\left(\hat{A} - \hat{A}^\dagger\right)|a_m\rangle = 0$$

which is still zero by definition of self-adjoint. Hence

$$\begin{aligned}\langle a_n|\hat{A}|a_m\rangle &= \langle a_n|\hat{A}^\dagger|a_m\rangle \\ &= \left(\langle a_m|\hat{A}|a_n\rangle\right)^*.\end{aligned}$$

These are again assumed to be eigenvectors, so:

$$\begin{aligned}a_m\langle a_n|a_m\rangle &= (a_n\langle a_m|a_n\rangle)^* \\ &= a_n^*\langle a_n|a_m\rangle\end{aligned}$$

and from [t**H1**]:

$$a_m\langle a_n|a_m\rangle = a_n\langle a_n|a_m\rangle.$$

Hence

$$(a_m - a_n)\langle a_n|a_m\rangle = 0.$$

Since $a_m \neq a_n$, the inner product must equal zero.

QED

4.5.4 Examples of Self-Adjoint Operators

It is a postulate of quantum mechanics that all observable quantities are represented by self-adjoint operators. Therefore examples ought to include:

- the Hamiltonian $\hat{H}$ (whose corresponding observable is the energy E)
- the position operator $\hat{x}$ (whose corresponding observable is the position x)
- the momentum operator $\hat{p}$ (whose corresponding observable is the momentum p).

However, whether these are really self-adjoint depends on boundary conditions, and is a subtle point we will return to.

4.5.5 Key Example of a Self-adjoint Operator: the Identity Operator

One operator that is guaranteed to exist in any inner product space is the identity operator $\mathbb{I}$.

Theorem [$t\mathbb{1}$1] (*Existence of* $\mathbb{1}$):

$$\forall \mathcal{I} \exists \mathbb{1} : \mathcal{I} \to \mathcal{I} | \forall |a\rangle \in \mathcal{I}, \mathbb{1}|a\rangle = |a\rangle.$$

Proof: simply observe that $|a\rangle$ is in $\mathcal{I}$ by definition!

Theorem [$t\mathbb{1}$2] (*Hermiticity of* $\mathbb{1}$):

$$\mathbb{1} = \mathbb{1}^\dagger.$$

Proof:

From [**H1**]:

$$\langle a|\mathbb{1}|b\rangle = \left(\langle b|\mathbb{1}^\dagger|a\rangle\right)^* . \tag{4.4}$$

Since

$$\mathbb{1}|b\rangle = |b\rangle$$

it follows that

$$\begin{aligned} \langle a|\mathbb{1}|b\rangle &= \langle a|b\rangle \\ &= \left(\langle b|a\rangle\right)^* \end{aligned}$$

where the second line follows from [**I1**]. Substituting into Equation 4.4 gives

$$\left(\langle b|a\rangle\right)^* = \left(\langle b|\mathbb{1}^\dagger|a\rangle\right)^*$$

so

$$\langle b|a\rangle = \langle b|\mathbb{1}^\dagger|a\rangle.$$

Since this is true for all $\langle b|$, it follows that

$$|a\rangle = \mathbb{1}^\dagger|a\rangle.$$

But we also have that

$$|a\rangle = \mathbb{I}|a\rangle.$$

Since this is true for all $|a\rangle$, it follows that $\mathbb{I} = \mathbb{I}^\dagger$. QED

5 Normed Vector Spaces

5.1 Definition of a Normed Vector Space

A *Normed Vector Space* $\mathcal{N}$ is a vector space $\mathcal{V}$ over a field $\mathcal{F}$ equipped with a *Norm* $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ obeying the following axioms.

$\forall |a\rangle, |b\rangle \in V;\ \alpha \in F:$

- **[N1]**: $\||a\rangle\| \geq 0$ (*non-negativity*)
- **[N2]**: $\||a\rangle\| = 0 \implies |a\rangle = |0\rangle$ (*positive definiteness*)
- **[N3]**: $\|\alpha|a\rangle\| = |\alpha| \, \||a\rangle\|$ (*absolute homogeneity*)
- **[N4]**: $\||a\rangle + |b\rangle\| \leq \||a\rangle\| + \||b\rangle\|$ (*subadditivity*, *aka* the *triangle inequality*).

Equivalently, a normed vector space is an ordered pair $\mathcal{N} = (\mathcal{V}, \|\cdot\|)$ obeying these axioms.

5.2 Examples of Normed Vector Spaces

Lemma [*l***I2**] implies that we can always define the norm in any inner product space:

For a vector $|a\rangle$ in an inner product space $\mathcal{I}$, the *Norm induced by the inner product* is defined to be

$$\||a\rangle\| \triangleq \sqrt{\langle a|a\rangle}$$

.

Hence, every inner product space is also a normed vector space: $\mathcal{I} \subseteq \mathcal{N}$. As a result, all the examples of inner product spaces serve as examples of normed vector spaces.

6 The Cauchy Schwarz Inequality

A very important corollary of the axioms of $\mathcal{I}$ is *The Cauchy-Schwarz inequality* (CS):

$$|\langle a|b\rangle| \leq \||a\rangle\| \, \||b\rangle\| \, .$$

Assuming the norm induced by the inner product, this takes the form:

$$|\langle a|b\rangle| \leq \sqrt{\langle a|a\rangle\langle b|b\rangle}.$$

Proof:

There are various proofs of this statement. I will provide one example here.

Define a function

$$p\left(t\right) \triangleq \|t\alpha|a\rangle + |b\rangle\|^2 \, .$$

Here,

$$p : \mathbb{R} \to \mathbb{R}$$

meaning the function takes $t \in \mathbb{R}$ as an argument and returns $p\left(t\right) \in \mathbb{R}$.

Define $\alpha \in \mathbb{C}$ as follows:

$$\begin{aligned} \alpha\langle a|b\rangle \triangleq |\langle a|b\rangle| \, , \langle a|b\rangle \neq 0 \\ \alpha \triangleq 1, \langle a|b\rangle = 0. \end{aligned}$$

Furthermore, we'll take the norm induced by the inner product. Therefore, expanding $p\left(t\right)$, we have

$$
\begin{aligned}
p(t) &= (t\alpha^* \langle a| + \langle b|)(t\alpha|a\rangle + |b\rangle) \\
&= t^2 |\alpha|^2 \langle a|a\rangle + \langle b|b\rangle + t(\alpha^* \langle a|b\rangle + \alpha \langle b|a\rangle) \\
&= t^2 |\alpha|^2 \||a\rangle\|^2 + \||b\rangle\|^2 + t(\alpha^* \langle a|b\rangle + \alpha \langle a|b\rangle) \\
&= t^2 |\alpha|^2 \||a\rangle\|^2 + \||b\rangle\|^2 + 2t\mathfrak{Re}(\alpha \langle a|b\rangle).
\end{aligned}
$$

Now use the definition of α to rewrite the first and last terms, noting that the definition implies $|\alpha| = 1$:

$$
p(t) = t^2 \||a\rangle\|^2 + \||b\rangle\|^2 + 2t |\langle a|b\rangle|.
$$

First let's check the case $|a\rangle = |0\rangle$. In this case, CS states

$$
|\langle 0|b\rangle| \leq \||0\rangle\| \||b\rangle\|.
$$

$$
\begin{aligned}
[\mathrm{l}\mathbf{I1}] &\implies \langle 0|b\rangle = 0 \\
[\mathbf{I3}] &\implies \||0\rangle\|
\end{aligned}
$$

therefore CS is trivially satisfied, with exact equality.

Next let's check $|a\rangle \neq |0\rangle$. In this case, $p \in \mathbb{P}_2$ (meaning p is a 2nd degree polynomial). However, since p is defined to be the square of a norm, [**I3**] implies that $p > 0$. Hence, p cannot change sign, and hence it has zero real roots. By the theory of quadratic equations, this means that the discriminant

$$
\Delta \triangleq 4\left(|\langle a|b\rangle|^2 - \||a\rangle\|^2 \||b\rangle\|^2\right) < 0.
$$

Hence,

$$
\begin{aligned}
|\langle a|b\rangle|^2 &< \||a\rangle\|^2 \||b\rangle\|^2 \\
|\langle a|b\rangle| &< \||a\rangle\| \||b\rangle\|
\end{aligned}
$$

in the case that $|a\rangle \neq |0\rangle$. Combining with the special case that $|a\rangle = |0\rangle$ gives the final result:

$$
|\langle a|b\rangle| \leq \||a\rangle\| \||b\rangle\|
$$

which is CS.

QED

6.1 Application of Cauchy Schwarz: The Heisenberg Uncertainty Principle

The proof proceeds as follows.

A. Define a few key objects used generally in quantum mechanics.

B. Define a few objects required for this specific proof.

C. Establish results $(i) - (vi)$ required for the proof.

D. Carry out the proof.

A. Define a few key objects used in quantum mechanics

Consider a Hermitian operator $\hat{A}$. Define the expectation value of $\hat{A}$ for the state $|\psi\rangle$ to be

$$\langle \hat{A} \rangle \triangleq \langle \psi | \hat{A} | \psi \rangle$$

and the 'uncertainty' in $\hat{A}$ to be its standard deviation:

$$\Delta A \triangleq \sqrt{\langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2}.$$

Finally, define the

$$\begin{aligned} \text{commutator:} \quad & \left[\hat{A}, \hat{B}\right] \triangleq \hat{A}\hat{B} - \hat{B}\hat{A} \\ \text{anticommutator:} \quad & \left\{\hat{A}, \hat{B}\right\} \triangleq \hat{A}\hat{B} + \hat{B}\hat{A}. \end{aligned}$$

B. Define a few objects required for this specific proof

Let

$$\delta\hat{A} \triangleq \hat{A} - \langle \hat{A} \rangle$$

such that

$$(\Delta A)^2 = \langle \left(\delta \hat{A}\right)^2 \rangle. \tag{6.1}$$

We will also define

$$\begin{aligned} |a\rangle &\triangleq \delta\hat{A}|\psi\rangle \\ |b\rangle &\triangleq \delta\hat{B}|\psi\rangle. \end{aligned}$$

C. Establish results $(i) - (vi)$ required for the proof.

$$(i) \qquad \delta\hat{A}\delta\hat{B} = \frac{1}{2}\left\{\delta\hat{A}, \delta\hat{B}\right\} + \frac{1}{2}\left[\delta\hat{A}, \delta\hat{B}\right]$$

as can be seen by expanding the right hand side;

$$(ii) \qquad \left[\delta\hat{A}, \delta\hat{B}\right] = \left[\hat{A}, \hat{B}\right]$$

similarly seen by expanding the right hand side.

$$(iii) \qquad \left(\left\{\delta\hat{A}, \delta\hat{B}\right\}\right)^{\dagger} = \left\{\delta\hat{A}, \delta\hat{B}\right\}$$

i.e. the anticommutator is Hermitian, and

$$(iv) \qquad \left(\left[\delta\hat{A}, \delta\hat{B}\right]\right)^{\dagger} = -\left[\delta\hat{A}, \delta\hat{B}\right]$$

i.e. the commutator is antihermitian.

Now we establish two important corollaries.

From (iii) it follows that the expectation value of the anticommutator is always purely real:

$$\begin{aligned}
\left\{\delta\hat{A}, \delta\hat{B}\right\} &= \left(\left\{\delta\hat{A}, \delta\hat{B}\right\}\right)^{\dagger} \\
\left\langle\left\{\delta\hat{A}, \delta\hat{B}\right\}\right\rangle &= \left\langle\left(\left\{\delta\hat{A}, \delta\hat{B}\right\}\right)^{\dagger}\right\rangle \\
\left\langle\left\{\delta\hat{A}, \delta\hat{B}\right\}\right\rangle &= \left\langle\left\{\delta\hat{A}, \delta\hat{B}\right\}\right\rangle^{*}
\end{aligned}$$

which gives us

$$(v) \qquad \left\langle\left\{\delta\hat{A}, \delta\hat{B}\right\}\right\rangle \in \mathbb{R}.$$

Similarly, from (iv), the expectation value of the commutator is always purely imaginary:

$$(vi) \qquad \left\langle\left[\delta\hat{A}, \delta\hat{B}\right]\right\rangle \in i\mathbb{R}.$$

D. Carry out the proof

From the definitions in B we have

$$
\begin{aligned}
\langle a|a\rangle &= \langle\psi|\delta\hat{A}\delta\hat{A}|\psi\rangle \\
&= \langle\psi|\left(\delta\hat{A}\right)^2|\psi\rangle \\
&= (\Delta A)^2
\end{aligned}
$$

where the last line follows from Equation 6.1.

Now invoke CS:

$$
\begin{gathered}
\||a\rangle\|^2\,\||b\rangle\|^2 \geq |\langle a|b\rangle|^2 \\
\downarrow \\
\langle a|a\rangle\langle b|b\rangle \geq |\langle a|b\rangle|^2 \\
\downarrow \\
(\Delta A)^2\,(\Delta B)^2 \geq \left|\langle\psi|\delta\hat{A}\delta\hat{B}|\psi\rangle\right|^2.
\end{gathered}
$$

Using (i) and (ii):

$$
(i) \implies (\Delta A)^2\,(\Delta B)^2 \geq \frac{1}{4}\left|\left\langle\left\{\delta\hat{A},\delta\hat{B}\right\}\right\rangle + \left\langle\left[\delta\hat{A},\delta\hat{B}\right]\right\rangle\right|^2
$$

and

$$
(ii) \implies (\Delta A)^2\,(\Delta B)^2 \geq \frac{1}{4}\left|\left\langle\left\{\delta\hat{A},\delta\hat{B}\right\}\right\rangle + \left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|^2.
$$

But (v) and (vi) tell us that the two terms within the modulus are purely real, and purely imaginary, respectively. Recalling that

$$
|x+iy|^2 = x^2+y^2,\ \{x,y\}\in\mathbb{R}
$$

leads us to conclude that

$$
(\Delta A)^2\,(\Delta B)^2 \geq \frac{1}{4}\left|\left\langle\left\{\delta\hat{A},\delta\hat{B}\right\}\right\rangle\right|^2 + \frac{1}{4}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|^2.
$$

But it is also the case that

$$
x^2+y^2 \geq y^2
$$

and so

$$\frac{1}{4}\left|\left\langle\left\{\delta\hat{A},\delta\hat{B}\right\}\right\rangle\right|^2+\frac{1}{4}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|^2\geq\frac{1}{4}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|^2.$$

Hence,

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

Square rooting gives the final result, *The Heisenberg Uncertainty principle*:

$$\Delta A\Delta B\geq\frac{1}{2}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|.$$

6.2 The Triangle Inequality

Using the axioms of inner product spaces we proved CS. Conversely, assuming CS we can prove [**N4**], the triangle inequality.

Proof:

Consider

$$\begin{aligned}\left\||x\rangle+|y\rangle\right\|^2&=\left(\langle x|+\langle y|\right)\left(|x\rangle+|y\rangle\right)\\&=\langle x|x\rangle+\langle y|y\rangle+\langle x|y\rangle+\langle y|x\rangle\\&=\langle x|x\rangle+\langle y|y\rangle+2\mathfrak{Re}\left(\langle x|y\rangle\right).\end{aligned}$$

Note that $\forall z\in\mathbb{C}$, $|z|\geq\mathfrak{Re}\left(z\right)$ (just write $z=x+iy$ and expand). Therefore

$$\left\||x\rangle+|y\rangle\right\|^2\leq\langle x|x\rangle+\langle y|y\rangle+2\left|\langle x|y\rangle\right|.$$

Now use CS, which says that

$$\left|\langle x|y\rangle\right|\leq\left\||x\rangle\right\|\left\||y\rangle\right\|$$

to give

$$\left\||x\rangle+|y\rangle\right\|^2\leq\langle x|x\rangle+\langle y|y\rangle+2\left\||x\rangle\right\|\left\||y\rangle\right\|.$$

Therefore

$$\||x\rangle + |y\rangle\|^2 \leq \||x\rangle\|^2 + \||y\rangle\|^2 + 2\,\||x\rangle\|\,\||y\rangle\|$$

and so

$$\||x\rangle + |y\rangle\|^2 \leq (\||x\rangle\| + \||y\rangle\|)^2$$

or

$$\||x\rangle + |y\rangle\| \leq \||x\rangle\| + \||y\rangle\|$$

which is the triangle inequality, [**N4**].

7 Metric Spaces

7.1 Definition of a Metric Space

A metric space $\mathcal{M}$ is a set M equipped with a map $d : M \times M \to \mathbb{R}$ (the *metric*). The metric satisfies the following axioms:

- **[M1]**: $d(x, y) \geq 0$, and $d(x, y) = 0 \Longleftrightarrow x = y$ (*positivity*)
- **[M2]**: $d(x, y) = d(y, x)$ (*symmetry*)
- **[M3]**: $d(x, z) \leq d(x, y) + d(y, z)$ (*subadditivity / triangle inequality*).

Equivalently, a metric space is an ordered pair $\mathcal{M} = (M, d)$ obeying these axioms.

Note: a metric space need not be a vector space, and the map (metric) need not be linear.

7.2 Examples of Metric Spaces

The norm is one example of a metric. Hence, all normed vector spaces $\mathcal{N}$ are metric spaces $\mathcal{M}$: $\mathcal{N} \subset \mathcal{M}$.

Specifically, some cases of interest are:

- $\mathcal{M} = ((M, \{+, \cdot\}), d)$: $M = \mathbb{R}$, $d(x, y) = |x - y|$ (real numbers, with their distance defined with the Euclidean metric in the usual way)
- $\mathcal{M} = ((M, \{+, \cdot\}), d)$: $M = V$, $d(|x\rangle, |y\rangle) = \||x\rangle - |y\rangle\|$ (a normed vector space with the metric induced by the norm).

8 Complete Spaces

8.1 Preamble: Equivalence

A relation $\sim$ on a set S is an *equivalence relation* iff $\forall a, b, c \in S$,

- **[E1]**: $a \sim a$ (*reflexivity*)
- **[E2]**: $a \sim b \iff b \sim a$ (*symmetry*)
- **[E3]**: $a \sim b$ & $b \sim c \implies a \sim c$ (*transitivity*).

We can supplement this with another useful definition, the *equivalence class* $[\cdot]$:

$$\forall a \in S, [a] \triangleq \{b \in S | a \sim b\}.$$

An important result is that an equivalence relation on a set S provides a partition of S into disjoint subsets called *equivalence classes*. That is, it allows you to divide the elements of S into subsets with no overlapping elements (equivalence classes).

8.2 Cauchy Completeness

Consider a sequence of rational numbers $x_n \in \mathbb{Q}$

$$x \triangleq (\, x_1,\, x_2,\, x_3,\, ...\,)$$

and a metric $d\,(x_m, x_n)$. The sequence x is *Cauchy* iff:

$$\forall \epsilon \in \mathbb{Q} > 0,\, \exists N, n, m \in \mathbb{N} | \forall n, m > N,\, d\,(x_m, x_n) < \epsilon.$$

That is, for any arbitrarily small ϵ, there exists some term N in the sequence beyond which *any* pair of terms $x_{m>N}$ and $x_{n>N}$ is separated by less than ϵ. This specifies a precise sense in which the sequence converges.

The sum of two Cauchy sequences is defined term-by-term:

$$x + y \triangleq (\, x_1 + y_1,\, x_2 + y_2,\, x_3 + y_3,\, ...\,)\,.$$

Two Cauchy sequences are equivalent under the metric d iff

$$x \sim y \iff \forall \epsilon \in \mathbb{Q} > 0,\, \exists N, n \in \mathbb{N} | \forall n > N,\, d\,(x_n, y_n) < \epsilon.$$

This defines a precise sense in which the sequences converge to the same limit (equivalently, their difference converges to zero). Here, $\sim$ is a formal equivalence relation obeying **E1-E3**. Hence, Cauchy sequences defined on a set S divide the set into disjoint equivalence classes.

8.2.1 Example: $\mathbb{Q} \hookrightarrow \mathbb{R}$

The set of rationals $\mathbb{Q}$ is 'incomplete'. For example, polynomials such as

$$x^2 - 2 = 0$$

have rational coefficients but no rational solutions.

Infinite series/sums of rationals can be defined which tend/evaluate to things which are not rational.

However, Cauchy sequences give a formal method of completing the rationals. The completion of a space is denoted $\hookrightarrow$. Consider the set X of all Cauchy sequences with elements in $\mathbb{Q}$. Choosing d to be the usual Euclidean metric provides a unique partition of X into equivalence classes. We define (!) this partition to be $\mathbb{R}$.

Hence we have the *definition of real numbers* ($\mathbb{R}$):

$\mathbb{R}$ is the set of all equivalence classes of Cauchy sequences in $\mathbb{Q}$ under the Euclidean metric.

For example, e is the equivalence class of the Cauchy sequence

$$e = (2,\, 2.7,\, 2.71,\, 2.718,\, 2.7182,\, ...)\,.$$

An equivalent sequence, also in this equivalence class, is defined by

$$x_{n>0} = (1 + 1/n)^n \,:\; (2,\, 9/4,\, 64/27,\, ...)$$

and another by

$$x_{n>0} = \sum_{i=0}^{n} \frac{1}{i!} \,:\; (2,\, 5/2,\, 8/3,\, ...)\,.$$

Note that any rational q can be represented as a Cauchy sequence simply by writing

$$q = (q, q, q, ...)$$

hence $\mathbb{Q} \subset \mathbb{R}$.

8.3 Definition of a Complete Metric Space

A metric space $\mathcal{M}$ is *complete* iff every Cauchy sequence of points in $\mathcal{M}$ converges to a point in $\mathcal{M}$.

9 Hilbert Spaces

9.1 Definition of a Hilbert Space

Finally we get to the most important definition of the course (you *will* need to learn this!!):

A Hilbert space is an inner product space that is also a complete metric space with respect to the norm induced by the inner product.

9.2 Examples of Hilbert Spaces

- The completion of any inner product space $\mathscr{I}$, using as the metric the norm induced by the inner product, is a Hilbert space: $\mathscr{I} \hookrightarrow \mathscr{H}$.
- All finite dimensional inner product spaces are Hilbert spaces. An example in quantum mechanics is the 2D complex inner product space describing spin-1/2.
- However, there are also infinite-dimensional Hilbert spaces, such as the space spanned by energy eigenstates of the particle in a box (1D quantum well), or particle on a ring.

9.3 Key Example of a Hilbert Space: Quantum Mechanics

One presentation of the axioms of quantum mechanics is as follows.

- **[QM1]** (*states*): States of a quantum system are represented by equivalence classes of unit vectors (*kets*) $[|\psi\rangle]$ in a Hilbert space $\mathscr{H}$, $[|\psi\rangle] \in \mathscr{H}$, under $|\psi\rangle \sim \exp(i\theta)|\psi\rangle$.
- **[QM2]** (*observables*):
  - **[QM2a]**: Observables are represented by self-adjoint linear maps (*operators*) $\hat{A} : \mathscr{H} \to \mathscr{H}$.
    * Corollary: by the spectral theorem, $\hat{A} = \int_{\mathbb{R}} \lambda \mathrm{d}\hat{E}_A(\lambda)$.
    * Corollary: if the spectrum is pure point, then $\hat{A} = \sum_n^{\dim \mathscr{H}} a_n |a_n\rangle\langle a_n|$.
  - **[QM2b]**: One such observable/operator pair is the energy/Hamiltonian $\hat{H}$.

- **[QM3]** (*measurement*):
  - **[QM3a]**: The possible results of a measurement of $\hat{A}$ on state $|\psi\rangle$ are in the spectrum of $\hat{A}$.
    - Corollary: if the spectrum is pure point, then the outcomes are a_n.
  - **[QM3b]** (*Born rule*): The probability of obtaining a result in the spectral set Δ is $\langle\psi|\hat{E}_A(\Delta)|\psi\rangle$.
    - Corollary: if the spectrum is pure point, then the probability is $|\langle a_n|\psi\rangle|^2$.
  - **[QM3c]**: Immediately after obtaining result a_n in a discrete spectrum, the system is in the corresponding eigenstate $|a_n\rangle$.
- **[QM4]** (*Dynamics*): In the absence of measurement, states evolve unitarily: $|\psi(t)\rangle = \exp\left(-i\hat{H}t/\hbar\right)|\psi(0)\rangle$.

9.4 Resolution of the Identity

9.4.1 Pre-amble: types of infinity

The number of elements in a set is called the cardinality of the set. We denote the cardinality of the integers to be $\aleph_0$ (aleph-null). It is also called countable infinity. This is known to be the smallest infinite cardinality. Interestingly, the cardinality of the rationals is also $\aleph_0$.

The cardinality of the reals, aka the *cardinality of the continuum*, is denoted $\mathfrak{c}$. Cantor's diagonal argument proves that $\mathfrak{c} = 2^{\aleph_0} > \aleph_0$, *i.e.* the cardinality of the continuum is strictly greater than the cardinality of the integers.

9.4.2 Separable Hilbert spaces

Definition: A Hilbert space is *separable* iff it admits an orthonormal basis of dimension at most $\aleph_0$.

N.B. *Non*-separable Hilbert spaces exist. For example, a chain of N spins-1/2 (the 1D Ising model) has Hilbert space dimension 2^N, and so an infinitely long chain has dimension $2^{\aleph_0} > \aleph_0$. But often we are safe in assuming our Hilbert space is separable, and separable spaces will be the focus here.

The statement that a separable Hilbert space is *complete* is equivalent to the statement that one can write the

Resolution of the identity:

given an orthonormal basis $\{|e_i\rangle\}$ for a separable Hilbert space $\mathcal{H}$,

$$\mathbb{I} = \sum_{i=1}^{\dim\mathcal{H}} |e_i\rangle\langle e_i|$$

where $\dim\mathcal{H}$ is the dimension of the Hilbert space (the number of linearly independent basis vectors $|e_i\rangle$).

To motivate this, consider the simple case of a 2D real inner product space $\mathcal{I} = (\mathbb{R}^2, \langle\cdot|\cdot\rangle)$, which, with the norm induced by the inner product, is a Hilbert space. Here, the dimension of the Hilbert space is finite (2), so it is separable.

Let

$$|v\rangle = v_1|e_1\rangle + v_2|e_2\rangle \tag{9.1}$$

where $\{|e_i\rangle\}$, $i \in [1, 2]$ form a complete orthonormal basis for $\mathcal{I}$. That is,

$$\langle e_i|e_j\rangle = \delta_{ij}$$

where we have defined *The Kronecker delta*:

$$\delta_{ij} \triangleq \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases} \tag{9.2}$$

Acting $\langle e_i|$ from the left on Equation 9.1 gives

$$v_i = \langle e_i|v\rangle$$

and so

$$|v\rangle = (\langle e_1|v\rangle)\,|e_1\rangle + (\langle e_2|v\rangle)\,|e_2\rangle$$

or

$$|v\rangle = |e_1\rangle\langle e_1|v\rangle + |e_2\rangle\langle e_2|v\rangle.$$

Factoring out the vector, we find

$$|v\rangle = (|e_1\rangle\langle e_1| + |e_2\rangle\langle e_2|)\,|v\rangle.$$

Since $|v\rangle$ is any vector in the space, the term in parentheses must be the identity operator. That is,

$$\mathbb{I} = |e_1\rangle\langle e_1| + |e_2\rangle\langle e_2|.$$

The same reasoning follows for any separable Hilbert space, defined to have a dimension that is at most $\aleph_0$.

10 Dirac vs Von Neumann

Separable Hilbert spaces have countable dimensions. But the positions of particles are quantified with real numbers, whose cardinality is $\mathfrak{c}$. This is not in itself a problem. To see why, consider the specific case of the infinite potential well:

$$V(x) = \begin{cases} 0, & 0 \leq x \leq L \\ \infty, & \text{otherwise} \end{cases}.$$

Another way to view this is as a free particle with fixed boundary conditions. Since the potential is zero within the well we have

$$\hat{H} = \frac{\hat{p}^2}{2m}$$

and hence the energy eigenstates should also be eigenstates of squared-momentum. Working in the position basis, the time independent Schroedinger equation reads

$$\hat{H}\phi_n(x) = \frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\phi_n(x) = E_n\phi_n(x)$$

and we find

$$E_n = \frac{\pi^2\hbar^2 n^2}{2mL^2}$$
$$\phi_n(x) = \frac{1}{\sqrt{L}}\sin\left(\frac{n\pi x}{L}\right).$$

Since the eigenenergies are non-degenerate, $\phi_n(x)$ form a complete orthonormal basis for the Hilbert space. Its dimension is $\aleph_0$ since the eigenstates are labelled by integers. Vitally, we can note that *any* (sensibly behaved) function of position $f(x)$, $x \in [0, L]$ can be decomposed into this basis:

$$f(x) = \sum_{n=1}^{\aleph_0} f_n\phi_n(x) \tag{10.1}$$

where

$$f_n \triangleq \int_0^L \phi_n^* (x) f (x) \, \mathrm{d}x. \tag{10.2}$$

So ***we can use a countable basis to describe functions of position, even though positions themselves have the cardinality of the continuum.*** This is not a problem or paradox. Dirac and Von Neumann agree on it, and indeed Von Neumann uses these results as the motivation for his own formalism.

The point of disagreement comes when Dirac introduces his famous notation. He says we can work in a basis-independent way by saying

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$$

and defining

$$\phi_n (x) \triangleq \langle x|\phi_n\rangle.$$

This is appealing, as we can rewrite Equation 10.1 in a basis-independent way (*i.e.* not in the position basis, or momentum basis, or any other basis) by similarly defining

$$\langle x|f\rangle = f (x)$$

to give

$$|f\rangle = \sum_{n=1}^{\aleph_0} f_n|\phi_n\rangle$$

where Equation 10.2 becomes

$$f_n = \langle\phi_n|f\rangle. \tag{10.3}$$

Combining these equations we find

$$|f\rangle = \sum_{n=1}^{\aleph_0} f_n|\phi_n\rangle = \sum_{n=1}^{\aleph_0}\langle\phi_n|f\rangle|\phi_n\rangle = \sum_{n=1}^{\aleph_0} |\phi_n\rangle\langle\phi_n|f\rangle$$

and since this holds for all $|f\rangle$ it follows that

$$\mathbb{I} = \sum_{n=1}^{\aleph_0} |\phi_n\rangle\langle\phi_n|. \tag{10.4}$$

The issue is that the notation suggests position eigenstates $|x\rangle$ are themselves elements of a Hilbert space. For example, comparing Equation 10.2 and Equation 10.3:

$$\begin{aligned}\langle\phi_n|f\rangle &= \int_0^L \phi_n^*(x) f(x)\,\mathrm{d}x \\ &= \int_0^L \langle\phi_n|x\rangle\langle x|f\rangle \mathrm{d}x \\ &= \langle\phi_n| \left(\int_0^L |x\rangle\langle x|\mathrm{d}x\right) |f\rangle\end{aligned}$$

and since this holds for all $\langle\phi_n|$ and all $|f\rangle$, it follows that

$$\mathbb{I} = \int_0^L |x\rangle\langle x|\mathrm{d}x. \tag{10.5}$$

By construction this resembles Equation 10.4. But Equation 10.5 assigns a basis element $|x\rangle$ to each point on the real interval $x \in [0, L]$, and hence the basis has dimension $\mathfrak{c} > \aleph_0$! With this, Von Neumann had a big problem, calling it a 'mathematical fiction'.

While it is tempting to suggest that these states simply exist in a non-separable Hilbert space, in fact they do not. Dirac recognised that they cannot live in a Hilbert space at all, and so rejected the idea that a Hilbert space was what we really use in quantum mechanics. Instead he formalised his orthonormal basis using

$$\langle x|y\rangle = \delta(x-y)$$

defining *The Dirac delta function*:

$$\delta(x-y) = \begin{cases} \infty, & x = y \\ 0, & x \neq y \end{cases}$$

normalised such that

$$\forall\epsilon > 0 \qquad \int_{y-\epsilon}^{y+\epsilon} \delta(x-y)\,\mathrm{d}x = 1.$$

This is to be thought of as a formal extension of the Kronecker delta, Equation 9.2. If we accept Dirac's notation, we can do neat tricks such as instantly rewrite into the momentum basis:

$$\tilde{f}(p) = \langle p|f\rangle$$

which naturally agrees with the usual Fourier transform:

$$\begin{aligned}
\tilde{f}(p) &= \langle p|f\rangle \\
&= \langle p|\mathbb{1}|f\rangle \\
&= \langle p|\int_0^L |x\rangle\langle x|\mathrm{d}x|f\rangle \\
&= \int_0^L \langle p|x\rangle\langle x|f\rangle \mathrm{d}x \\
&= \int_0^L \exp\left(-ipx/\hbar\right) f(x)\,\mathrm{d}x
\end{aligned}$$

where the momentum eigenstate is a plane wave:

$$\langle x|p\rangle = \exp\left(ipx/\hbar\right).$$

10.1 Square-summable functions

A function $f : [a, b] \in \mathbb{R} \to \mathbb{C}$ is *square summable* (aka *square integrable*) iff

$$\int_a^b |f(x)|^2 \,\mathrm{d}x < \infty. \tag{10.6}$$

Here, $a < b$, but neither need be finite. Square summable functions form an inner product space, conventionally denoted $\mathcal{L}^2 = \left(L^2, \langle\cdot|\cdot\rangle_2\right)$, where L^2 (pronounced *L-two*) is the set of square summable functions under the norm induced by the inner product $\langle\cdot|\cdot\rangle_2$. It can be shown that any such space is complete. Hence, $\mathcal{L}^2$ is a Hilbert space[1].

Following Dirac, we can rewrite Equation 10.6 as

[1]Prof. Robbins comments: "provided integration and square-integrability are understood in the sense of Lebesgue. If integration is in the sense of Riemann, then the space of square-integrable functions is not complete."

$$\int_a^b \langle f|x\rangle\langle x|f\rangle \mathrm{d}x < \infty$$

$$\downarrow$$

$$\langle f| \left(\int_a^b |x\rangle\langle x| \mathrm{d}x \right) |f\rangle < \infty.$$

Therefore square-summability in Dirac notation reads

$$\langle f|f\rangle < \infty.$$

But in quantum mechanics states are normalised:

$$\langle f|f\rangle = 1$$

so this is automatically true. The point is that you would be assuming square-summability naturally as a physicist.

We can similarly define $\mathscr{L}^p$ spaces, $p > 2$, as sets of p-summable functions with a suitably defined metric:

$$\|f\|_p \triangleq \int_a^b |f(x)|^p \, \mathrm{d}x < \infty.$$

However, $p = 2$ is the only case in which the metric $\|\cdot\|_p$ is compatible with an inner product. Hence, $\mathscr{L}^2$ is the only $\mathscr{L}^p$ space which is also a Hilbert space.

11 Fourier Series

Consider functions defined on the real interval $x \in [-L, L)$. The aim is to decompose them into the following components:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{L}\right). \tag{11.1}$$

Recall from first year that

$$\begin{aligned} a_0 &= \frac{1}{L}\int_{-L}^{L} f(x)\, \mathrm{d}x \\ a_{n>0} &= \frac{1}{L}\int_{-L}^{L} f(x) \cos\left(\frac{n\pi x}{L}\right) \mathrm{d}x \\ b_n &= \frac{1}{L}\int_{-L}^{L} f(x) \sin\left(\frac{n\pi x}{L}\right) \mathrm{d}x. \end{aligned}$$

11.1 Fourier Series in Dirac notation

Define some abstract states in an $\aleph_0$-dimensional Hilbert space $|0\rangle, |c_n\rangle, |s_n\rangle \in \mathscr{H}$ by their projections into the position basis:

$$\begin{aligned} \langle x|0\rangle &\triangleq \frac{1}{\sqrt{2L}} \\ \langle x|c_n\rangle &\triangleq \frac{1}{\sqrt{L}} \cos\left(\frac{n\pi x}{L}\right) \\ \langle x|s_n\rangle &\triangleq \frac{1}{\sqrt{L}} \sin\left(\frac{n\pi x}{L}\right). \end{aligned}$$

In this abstract Hilbert space, Equation 11.1 reads

$$|f\rangle = \frac{a_0}{2}|0\rangle + \sum_{n=1}^{\infty} a_n |c_n\rangle + \sum_{n=1}^{\infty} b_n |s_n\rangle$$

where

$$\begin{aligned} a_0/2 &= \langle 0|f\rangle \\ a_{n>0} &= \langle c_n|f\rangle \\ b_n &= \langle s_n|f\rangle. \end{aligned}$$

Implicit in this construction is the fact that $\{|0\rangle, |c_n\rangle, |s_n\rangle\}$ form a complete orthonormal basis for $\mathscr{H}$. That is,

$$\begin{aligned} \langle 0|c_n\rangle = \langle 0|s_n\rangle &= \langle c_n|s_n\rangle = 0 \\ \langle 0|0\rangle &= 1 \\ \langle c_n|c_m\rangle = \langle s_n|s_m\rangle &= \delta_{nm}. \end{aligned}$$

Each of these can be shown using a resolution of the identity into the position basis. *E.g.*

$$\begin{aligned} \langle 0|0\rangle &= \langle 0|\mathbb{1}|0\rangle \\ &= \langle 0| \left(\int_{-L}^{L} \mathrm{d}x |x\rangle\langle x| \right) |0\rangle \\ &= \int_{-L}^{L} \mathrm{d}x \langle 0|x\rangle\langle x|0\rangle \\ &= \int_{-L}^{L} \mathrm{d}x \frac{1}{\sqrt{2L}} \frac{1}{\sqrt{2L}} \\ 1 &\checkmark \end{aligned}$$

Completeness is (I'm told) equivalent to the statement that all basis vectors other than $|0\rangle$ are orthogonal to $|0\rangle$.

11.2 QM example: the infinite potential well

Consider a quantum particle in a potential

$$V(x) = \begin{cases} 0, & 0 \leq x \leq L \\ \infty, & \text{otherwise} \end{cases}$$

(the infinite potential well, or particle in a box). The task is to find the energy eigenstates defined by

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$$

where

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}$$
$$\hat{V} = \int_{-\infty}^{\infty} \mathrm{d}x V(x) |x\rangle\langle x|.$$

In the position basis this reads

$$\frac{-\hbar^2}{2m}\frac{\partial^2 \phi_n(x)}{\partial x^2} + V(x)\phi_n(x) = E_n \phi_n(x)$$

where

$$\phi_n(x) \triangleq \langle x|\phi_n\rangle.$$

The solutions are

$$\phi_n(x) = \frac{1}{\sqrt{L}} \sin\left(\frac{n\pi x}{L}\right)$$

with eigenenergies

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2mL^2}.$$

Note that

$$|\phi_n\rangle = |s_n\rangle$$

as defined above. With the boundary conditions $x \in [0, L]$ it follows that $a_0 = b_n = 0$. Since $|s_n\rangle$ are the normalised eigenstates of the Hermitian operator $\hat{H}$, with non-degenerate eigenvalues E_n, they form a complete orthonormal basis for $\mathcal{H}$. That is, any function $f(x)$ on the interval $x \in [0, L]$ can be written as a sum of $\langle x|s_n\rangle$.

12 Generalisation to Orthogonal Polynomials

In the proceeding chapters we will see various orthogonal polynomials. Each lives in an $\mathcal{L}^2$ space, which is always a Hilbert space. We must define a domain on which the polynomials are defined (an interval of the real line), say $[a, b]$, and a weight function $w(x)$ for inner products, such that

$$\langle f|g\rangle \triangleq \int_a^b f^*(x)\, g(x)\, w(x)\, \mathrm{d}x.$$

The next few sections look at some specific examples you will encounter elsewhere.

13 Legendre Polynomials

13.1 Definition of Legendre Polynomials

Legendre polynomials $P_n(x)$ are defined over the Hilbert space $\mathcal{L}^2[-1,1]$ (an $\mathcal{L}^2-$space defined over the interval $[-1,1] \in \mathbb{R}$) with weight function $w(x) = 1$. This is typically denoted as the ordered pair $\mathcal{H} = (\mathcal{L}^2[-1,1], w(x) = 1)$. This means that the inner product between two functions is

$$\langle f|g\rangle \triangleq \int_{-1}^{1} f(x)^* \, g(x) \, w(x) \, \mathrm{d}x$$

where $w(x) = 1$. Hence, it is the usual inner product for functions.

- The Legendre polynomials are solutions to Legendre's differential equation

$$(1 - x^2) \, y'' - 2xy' + n(n+1) \, y = 0.$$

- First few terms:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}\left(3x^2 - 1\right) \end{aligned}$$

- Rodrigues' formula:

$$P_n(x) = \frac{1}{2^n n!} \mathrm{d}_x^n \left(x^2 - 1\right)^n$$

- Recursion formula (Bonnet's formula): given P_0, P_1,

$$(n+1) \, P_{n+1}(x) = (2n+1) \, x P_n(x) - n P_{n-1}(x)$$

- Generating function

$$\left(1-2xt+t^2\right)^{-1/2}=\sum_{n=0}^{\infty}P_n\left(x\right)t^n$$

meaning

$$P_n\left(x\right)=\left.\partial_t^n\left(1-2xt+t^2\right)^{-1/2}\right|_{t=0}$$

- Contour integral expression

$$P_n\left(z\right)=\frac{1}{2\pi i}\oint_C\left(1-2tz+t^2\right)^{-1/2}t^{-n-1}\mathrm{d}t$$

where C encircles the origin anticlockwise.

13.2 Orthogonality & Normalisation

$$\int_{-1}^{1}P_n\left(x\right)P_m\left(x\right)\mathrm{d}x=\frac{2}{2n+1}\delta_{nm}.$$

13.3 Completeness

The Legendre Polynomials form an orthonormal basis for functions on the interval $[-1,1]$. To prove that they form a *complete* orthonormal basis is tricky. The proof is not covered here. The statement, though, is as follows.

Given any piecewise continuous function $f\left(x\right)$ with at most finitely many discontinuities on the interval $[-1,1]$, the sequence of partial sums

$$f_n\left(x\right)=\sum_{l=0}^{n}a_lP_l\left(x\right)$$

converges to $f\left(x\right)$ in the limit $n\to\infty$. Here,

$$a_l=\frac{2l+1}{2}\int_{-1}^{1}f\left(x\right)P_l\left(x\right)\mathrm{d}x.$$

Another way to state this is

$$\sum_{l=0}^{\aleph_0} \frac{2l+1}{2} P_l(x) P_l(y) = \delta(x-y)$$

with $x, y \in [-1, 1]$. This is a statement that we can resolve the identity into the basis of $|P_l\rangle$. Hence, The Legendre Polynomials form a complete basis for the Hilbert space $\mathcal{H} = (\mathcal{L}^2[-1,1], w(x) = 1)$.

13.4 QM example: spherical potential, angular part

If you are interested to see where Legendre Polynomials appear in quantum mechanics, please see my notes on that subject [https://www.felixflicker.com/files/PX2132/PX2132_notes_2022.pdf].

14 (Generalised) Laguerre Polynomials

14.1 Definition of Generalised Laguerre Polynomials

The generalised Laguerre Polynomials $L_n^k(x)$ (the Laguerre Polynomials have $k = 0$) are defined over the Hilbert space $\mathcal{L}^2[0, \infty)$ with weight function $w(x) = x^k \exp(-x)$. Hence, the inner product is now defined as

$$\langle f|g\rangle \triangleq \int_0^\infty f(x)^* g(x) x^k \exp(-x) \,\mathrm{d}x. \tag{14.1}$$

- They solve (the generalised) Laguerre's differential equation

$$xy'' + (k + 1 - x) y' + ny = 0$$

- First few terms:

$$\begin{aligned} L_0^k(x) &= 1 \\ L_1^k(x) &= k + 1 - x \\ L_2^k(x) &= \frac{1}{2}\left(x^2 - 2(k+2)x + (k+1)(k+2)\right) \end{aligned}$$

- Rodrigues' formula:

$$L_n^k(x) = \frac{x^{-k}}{n!}(\mathrm{d}_x - 1)^n x^{n+k}$$

- Recursion formula: given the first two terms, for $n \geq 2$:

$$L_{n+1}^k(x) = \frac{(2n + 1 + k - x) L_n^k(x) - (n + k) L_{n-1}^k(x)}{n + 1}$$

- Generating function

$$\frac{\exp\left(-tx/\left(1-t\right)\right)}{\left(1-t\right)^{k+1}} = \sum_{n=0}^{\infty} t^n L_n^k\left(x\right)$$

meaning

$$L_n^k\left(x\right) = \partial_t^n \frac{\exp\left(-tx/\left(1-t\right)\right)}{\left(1-t\right)^{k+1}}\Bigg|_{t=0}.$$

- Contour integral expression

$$L_n^k\left(x\right) = \frac{1}{2\pi i}\oint_C \frac{\exp\left(-tx/\left(1-t\right)\right)}{\left(1-t\right)^{k+1} t^{n+1}}\mathrm{d}t$$

where C encircles the origin anticlockwise but omits the essential singularity at 1.

14.2 Orthogonality, Normalisation and Completeness

The Laguerre polynomials are orthonormal with respect to the inner product in Equation 14.1:

$$\int_0^\infty L_n^k\left(x\right) L_m^k\left(x\right) x^k \exp\left(-x\right) \mathrm{d}x = \frac{(n+k)!}{n!}\delta_{nm}.$$

The functions

$$\varphi_n^k\left(x\right) \triangleq x^{k/2}\exp\left(-x/2\right) L_n^k\left(x\right)$$

form a complete basis for the Hilbert space $\mathcal{L}^2[0,\infty)$.

14.3 QM example: the Hydrogen atom

If you are interested to see where Laguerre polynomials appear in quantum mechanics, please see my notes on that subject [https://www.felixflicker.com/files/PX2132/PX2132_notes_2022.pdf].

15 Hermite Polynomials

15.1 Definition of Hermite Polynomials

The Hermite polynomials $H_n(x)$ are defined over the Hilbert space $\mathcal{L}^2(-\infty, \infty)$ with weight function $w(x) = \exp(-x^2)$.

- Hermite's differential equation

$$y'' + (2n + 1 - x^2)\, y = 0$$

is solved by the *Hermite functions*

$$\psi_n(x) = \left(2^n n! \sqrt{\pi}\right)^{-1/2} \exp\left(-x^2/2\right) H_n(x)$$

where $H_n(x)$ are the Hermite Polynomials.

- First few terms:

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 \end{aligned}$$

- Rodrigues' formula

$$H_n(x) = \left(2x - \mathrm{d}_x\right)^n \cdot 1$$

by which it is meant that the operator in parentheses acts on 1, *e.g.*

$$\begin{aligned} H_2(x) &= \left(2x - \mathrm{d}_x\right)^2 \cdot 1 \\ &= \left(2x - \mathrm{d}_x\right)\left(2x - \mathrm{d}_x\right) \cdot 1 \\ &= \left(4x^2 - 2x\mathrm{d}_x - \mathrm{d}_x(2x) - \mathrm{d}_x^2\right) \cdot 1 \\ &\downarrow \text{(chain rule on 3rd term)} \\ &= \left(4x^2 - 2x\mathrm{d}_x - 2 - 2x\mathrm{d}_x - \mathrm{d}_x^2\right) \cdot 1 \\ &= 4x^2 - 2. \end{aligned}$$

- Recursion formula: given $H_0(x)$,

$$H_{n+1}(x) = 2xH_n(x) - H_n'(x)$$

- Generating function

$$\exp\left(2xt - t^2\right) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x)\, t^n$$

meaning

$$H_n(x) = \partial_t^n \exp\left(2xt - t^2\right)\Big|_{t=0}.$$

- Contour integral expression

$$H_n(z) = \frac{n!}{2\pi i} \oint_C \exp\left(-t^2 + 2tz\right) t^{-n-1} \mathrm{d}t$$

where C encircles the origin anticlockwise.

15.2 Orthogonality, Normalisation and Completeness

The Hermite functions are orthogonal and normalised:

$$\int_{-\infty}^{\infty} \psi_n(x)\, \psi_m(x)\, \mathrm{d}x = \delta_{nm}.$$

They are also complete:

$$\sum_{n=0}^{\infty} \psi_n(x)\, \psi_n(y) = \delta(x - y).$$

Hence, they form a complete orthonormal basis for the Hilbert space $\mathcal{L}^2(\mathbb{R})$.

15.3 QM example: Quantum Harmonic Oscillator

A very important problem in quantum mechanics is the quantum harmonic oscillator:

$$\hat{V} = \frac{1}{2} m\omega^2 \hat{x}^2.$$

In the position basis, the time independent Schroedinger equation reads

$$-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2 \phi_n(x)}{\mathrm{d}x^2} + \frac{1}{2} m\omega^2 x^2 \phi_n(x) = E_n \phi_n(x).$$

It is convenient to rescale using $x = \alpha y$:

$$-\frac{\hbar^2}{2m\alpha^2} \frac{\mathrm{d}^2 \phi_n(y)}{\mathrm{d}y^2} + \frac{1}{2} m\omega^2 \alpha^2 y^2 \phi_n(y) = E_n \phi_n(y)$$

and selecting

$$\alpha^2 = \frac{\hbar}{m\omega}$$

gives

$$\frac{1}{2} \left(-\frac{\mathrm{d}^2 \phi_n(y)}{\mathrm{d}y^2} + y^2 \phi_n(y) \right) = \epsilon_n \phi_n(y) \tag{15.1}$$

where

$$\epsilon_n \triangleq \frac{E_n}{\hbar\omega}.$$

Equation 15.1 is a second order ODE. Substitute

$$\phi_n(y) = H_n(y) \exp\left(-\frac{y^2}{2} \right)$$

to reduce Equation 15.1 to

$$H_n(y)'' - 2y H_n(y)' + (2\epsilon - 1) H_n(y) = 0.$$

This is Hermite's equation. It can be solved with Frobenius series to yield $H_n(y)$, the Hermite polynomials for $n \geq 1$:

$$H_n(y) = (-1)^n \exp\left(y^2\right) \frac{\mathrm{d}^n}{\mathrm{d}y^n} \exp\left(-y^2\right)$$

(where $H_0 = 1$) with energy eigenvalues

$$\epsilon_n = n + 1/2$$

for integer $n \geq 0$. Returning to the original scaling we have

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right).$$

16 Completeness of Polynomials

We can rephrase all the cases of orthogonal polynomials above, including Fourier series, more directly into Dirac notation. The underlying principle is the following.

Let $\{|f_n\rangle\}$ be orthogonal on the interval $[a, b]$ with respect to the weight function $w(x)$:

$$\langle f_n | f_m \rangle = \int_a^b f_n(x)^* f_m(x) w(x) \, \mathrm{d}x \propto \delta_{nm}.$$

NB the existence (non-infiniteness) of this norm defines these spaces to be $L^2([a, b], w(x) \, \mathrm{d}x)$, and therefore Hilbert spaces:

$$\langle f_n | f_m \rangle = \int_a^b f_n(x)^* f_m(x) w(x) \, \mathrm{d}x < \infty.$$

Then we can always decompose any function $c(x) \in L^2([a, b], w(x) \, \mathrm{d}x)$ into the $\{|f_n\rangle\}$ basis:

$$c(x) = \sum_n c_n f_n(x)$$

with

$$c_n = \frac{\langle f_n | c \rangle}{\langle f_n | f_n \rangle} = \frac{\int_a^b f_n^*(x) c(x) w(x) \, \mathrm{d}x}{\int_a^b f_n^*(x) f_n(x) w(x) \, \mathrm{d}x}.$$

This is the main use of orthogonality – it allows a formal extension of vector decomposition to the case of functions.

The statement of completeness is equivalent to a statement of the existence of a resolution of the identity, as follows:

$$
\begin{aligned}
|c\rangle &= \sum_n c_n |f_n\rangle \\
&= \sum_n \left(\frac{\langle f_n | c\rangle}{\langle f_n | f_n\rangle} \right) |f_n\rangle \\
&= \sum_n \frac{|f_n\rangle\langle f_n|}{\langle f_n | f_n\rangle} |c\rangle
\end{aligned}
$$

and therefore

$$
\mathbb{I} = \sum_n \frac{|f_n\rangle\langle f_n|}{\langle f_n | f_n\rangle}.
$$

Assuming the basis is normalised (always possible), $\langle f_n | f_n\rangle = 1$, and

$$
\mathbb{I} = \sum_n |f_n\rangle\langle f_n|.
$$

The basis $\{|f_n\rangle\}$ is said to be **complete** if any of the equivalent statements is true:

- $\hat{\mathbb{I}} = \sum_n |f_n\rangle\langle f_n|$ (the identity can be decomposed into the basis)
- All functions $f_n(x) \in L^2([a,b], w(x)\,\mathrm{d}x)$ are orthogonal to the zero function
- The zero function is the only function orthogonal to all $f_n(x) \in L^2([a,b], w(x)\,\mathrm{d}x)$.

In general, proving completeness is difficult and subtle. Recall that **QM3**, the third axiom of quantum mechanics, has to *postulate* completeness!

17 Sturm Liouville Theory

It is natural to wonder why so many different orthogonal polynomials exist, with such similar underlying properties. The fact that they live in Hilbert spaces is a characterisation of this similarity, rather than a proof that it must be the case. The underlying connection is provided by Sturm Liouville theory. It turns out that all the orthogonal polynomials we have considered are special cases of this more general form.

17.1 The Regular Sturm Liouville Problem statement

Given functions $p(x)$, $q(x)$, the weight function $w(x)$, a real domain $x \in [a, b]$, and 'separated' boundary conditions, the *regular Sturm Liouville problem* is the second order linear ODE

$$\begin{aligned} (p(x)\, y'(x))' + q(x)\, y(x) &= -\lambda w(x)\, y(x)\,, \quad x \in [a, b] \\ c_1 y(a) + c_2 y'(a) &= 0 \quad \text{(boundary condition 1)} \\ d_1 y(b) + d_2 y'(b) &= 0 \quad \text{(boundary condition 2)} \end{aligned} \tag{17.1}$$

where:

(i)

$$\begin{aligned} (c_1, c_2) &\neq (0, 0) \\ (d_1, d_2) &\neq (0, 0) \end{aligned}$$

(ii) p, p', q, and w are continuous on $[a, b]$

(iii) p and w are positive definite on $[a, b]$.

Together, these conditions guarantee that solutions to Equation 17.1 exist (according to the standard theory of linear ODEs).

A solution to the regular Sturm Liouville problem is defined to be a non-trivial, continuously differentiable function $y(x)$ which solves Equation 17.1 on the interval (a, b), and which matches the boundary conditions. Such solutions are called eigenfunctions. These will only exist for certain values of λ, which are called the eigenvalues.

It can be shown that the Hermite, Legendre, and Laguerre polynomials are special cases of the Sturm Liouville problem, as is Fourier theory.

17.2 Example: simple harmonic oscillator

Consider the simple harmonic oscillator

$$\begin{aligned} y'' + \lambda y = 0, &\quad x \in [0, \pi] \\ y(0) = 0 & \\ y(\pi) = 0. & \end{aligned}$$

The solutions satisfying the boundary conditions are

$$y_n(x) = A \sin(nx), \quad n \in \mathbb{N}$$

with the eigenvalues

$$\lambda_n = n^2.$$

17.3 Properties of Sturm Liouville solutions

17.3.1 Real Eigenvalues

The eigenvalues of the regular Sturm Liouville problem are real:

$$\lambda_n \in \mathbb{R}.$$

17.3.2 Distinct Eigenvalues

The eigenvalues of the regular Sturm Liouville problem are non-degenerate:

$$\lambda_1 < \lambda_2 < \lambda_3 \ldots$$

17.3.3 Infinite number of eigenvalues

There exists a countably infinite number of eigenvalues: $n \in \aleph_0$.

17.3.4 Infinite extent of eigenvalues

The largest eigenvalue is infinitely large:

$$\lim_{n \to \aleph_0} \lambda_n = \infty.$$

17.3.5 Uniqueness of (normalised) Eigenfunctions

The eigenfunction $y_n(x)$ corresponding to eigenvalue λ_n is unique, up to a multiplicative constant. Equivalently, the normalised eigenfunctions are unique.

17.3.6 Number of zeroes

$y_n(x)$ has precisely $n-1$ zeroes in the interval $[a, b]$.

17.3.7 Orthogonal eigenfunctions

Distinct eigenfunctions are orthogonal:

$\langle y_n | y_m \rangle = \int_a^b y_n^* (x)\, y_m(x)\, w(x)\, \mathrm{d}x = 0$ if $n \neq m$.

17.3.8 Completeness

The normalised eigenfunctions $y_n(x)$ form a complete orthonormal basis for the Hilbert space $\mathcal{L}^2([a, b], w(x)\, \mathrm{d}x)$:

$$\langle y_n | y_m \rangle = \int_a^b y_n^* (x)\, y_m(x)\, w(x)\, \mathrm{d}x = \delta_{nm}.$$

Equivalently, given any function $c(x)$ in the interval, we can write

$$c(x) = \sum_{n=1}^{\infty} c_n y_n(x) \tag{17.2}$$

where

$$c_n = \frac{\langle y_n | c \rangle}{\langle y_n | y_n \rangle} = \frac{\int_a^b y_n^* (x)\, c(x)\, w(x)\, \mathrm{d}x}{\int_a^b y_n^* (x)\, y_n(x)\, w(x)\, \mathrm{d}x}.$$

The discrete sum in Equation 17.2 is at first sight remarkable, as nothing about the setup of the Sturm Liouville problem appeared to involve discreteness. At this point in the course it is probably expected! It derives from the constraints of the boundary conditions.

17.4 The Sturm Liouville Operator

One way to understand these properties is to notice that the Sturm Liouville problem can be written in terms of a Hermitian operator.

Define the differential operator

$$\hat{L} \triangleq -\frac{1}{w(x)}\left(\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}}{\mathrm{d}x}\right)+q(x)\right).$$

Then the Sturm Liouville problem reduces to

$$\hat{L}y_n(x)=\lambda_n y_n(x)$$

(with appropriate boundary conditions, and other conditions as specified in Section 1). Various properties outlined in Section 1.2 then follow immediately from the fact that $\hat{L}$ is Hermitian: for example, that the eigenvalues are real, and that distinct eigenfunctions are orthogonal.

17.5 Proof of Hermiticity of the Sturm Liouville Operator $\hat{L}$

We are required to show that

$$\left(\hat{L}|f\rangle\right)^{\dagger}|g\rangle=\langle f|\left(\hat{L}|g\rangle\right)$$

or equivalently

$$\int_a^b\left(\hat{L}f(x)\right)^{*}g(x)\,w(x)\,\mathrm{d}x=\int_a^b f^{*}(x)\left(\hat{L}g(x)\right)w(x)\,\mathrm{d}x.$$

For simplicity, let's consider the case $w(x)=1$. (In fact this is without loss of generality, as we can incorporate $w(x)$ into $y(x)$.)

Inserting the definition of $\hat{L}$ gives

$$\begin{aligned}\int_a^b \left(\hat{L}f(x)\right)^* g(x)\,\mathrm{d}x &= \int_a^b \left(-\left(\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\,f'(x)\right] + q(x)\,f(x)\right)\right)^* g(x)\,\mathrm{d}x \\ &= -\int_a^b \left(\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\,f'^*(x)\right] + q(x)\,f^*(x)\right) g(x)\,\mathrm{d}x \\ &= -\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\,f'^*(x)\right] g(x)\,\mathrm{d}x - \int_a^b q(x)\,f^*(x)\,g(x)\,\mathrm{d}x\end{aligned} \tag{17.3}$$

where we used the fact that p and q are specified to be real on the interval. Now integrate the first term of Equation 17.3 by parts:

$$\int_a^b \left(\hat{L}f(x)\right)^* g(x)\,\mathrm{d}x = -\left[p(x)\,f'^*(x)\,g(x)\right]_a^b + \int_a^b p(x)\,f'^*(x)\,g'(x)\,\mathrm{d}x - \int_a^b q(x)\,f^*(x)\,g(x)\,\mathrm{d}x$$

and integrate the second term by parts:

$$\int_a^b \left(\hat{L}f(x)\right)^* g(x)\,\mathrm{d}x = -\left[p(x)\,f'^*(x)\,g(x)\right]_a^b + \left[p(x)\,f^*(x)\,g'(x)\right]_a^b - \int_a^b f^*(x)\,\mathrm{d}_x\left[p(x)\,g'(x)\right]\mathrm{d}x - \int_a^b q(x)\,f^*(x)$$

and combine terms:

$$\begin{aligned}\int_a^b \left(\hat{L}f(x)\right)^* g(x)\,\mathrm{d}x &= \left[p(x)\left(f^*(x)\,g'(x) - f'^*(x)\,g(x)\right)\right]_a^b - \int_a^b f^*(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\,g'(x)\right] + q(x)\,g(x)\right)\mathrm{d}x \\ &= \left[p(x)\left(f^*(x)\,g'(x) - f'^*(x)\,g(x)\right)\right]_a^b + \int_a^b f^*(x)\left(\hat{L}g(x)\right)\mathrm{d}x.\end{aligned}$$

Therefore the operator $\hat{L}$ is Hermitian provided the boundary term vanishes:

$$\left[p(x)\left(f^*(x)\,g'(x) - f'^*(x)\,g(x)\right)\right]_a^b = 0.$$

This can be guaranteed by mandating that all functions under consideration obey the separated boundary conditions

$$\begin{aligned}c_1 y(a) + c_2 y'(a) &= 0 \\ d_1 y(b) + d_2 y'(b) &= 0\end{aligned} \tag{17.4}$$

where

$$\begin{aligned}(c_1, c_2) &\neq (0,0)\\ (d_1, d_2) &\neq (0,0)\,.\end{aligned}$$

To check this, note that Equation 17.4 can be rewritten

$$\begin{aligned}y'(a) &= -\frac{c_2}{c_1}y(a)\\ y'(b) &= -\frac{d_2}{d_1}y(b)\end{aligned}$$

(which is well defined because of the constraint that both terms are not simultaneously zero), and so the boundary term becomes

$$\begin{aligned}&\left[p(b)\left(f^*(b)\,g'(b) - f'^*(b)\,g(b)\right)\right] - \left[p(a)\left(f^*(a)\,g'(a) - f'^*(a)\,g(a)\right)\right]\\ =&\left[\frac{d_2}{d_1}p(b)\left(-f^*(b)\,g(b) + f^*(b)\,g(b)\right)\right] - \left[\frac{c_2}{c_1}p(a)\left(f^*(a)\,g(a) - f^*(a)\,g(a)\right)\right]\end{aligned}$$

and so each boundary vanishes separately.

If $p(a) = p(b)$, one can instead guarantee periodic solutions

$$\begin{aligned}f(a) &= f(b)\\ f'(a) &= f'(b)\,.\end{aligned}$$

18 Postscript

The concept of a Hilbert space was introduced in an attempt to formalise the mathematics which physicists were inventing in order to explain the bizarre behaviours on subatomic scales. After all you've seen in this course, does it do so? Not really!

This was known as soon as the theory was formulated. Dirac already wrote in his 1930 textbook:

> *The bra and ket vectors that we now use form a more general space than a Hilbert space.*

Von Neumann, similarly, rejected the concept of a Hilbert space as a description of QM by 1935.

The failure is severe. Essentially every operator you are likely to write down in quantum mechanics is a function of position and momentum (and possibly spin). On the other hand, the first postulate of quantum mechanics says that states live in a Hilbert space. Yet the eigenstates of position do not live in a Hilbert space!

The issues are already perfectly evident from consideration of the two simplest problems in quantum mechanics.

18.1 Plane waves

Let us return briefly to the distinction between bounded Hermitian operators and (potentially unbounded) self-adjoint operators in Section 4.4. The absolute simplest quantum problem has $V(x) = 0$, $x \in (-\infty, \infty)$ The solutions to the time independent Schroedinger equation

$$\hat{H}|p\rangle = \frac{\hat{p}^2}{2m}|p\rangle = E_n|p\rangle$$

are eigenstates of the momentum operator:

$$\hat{p}|p\rangle = p|p\rangle$$

where

$$\langle x|p\rangle \propto \exp\left(ipx/\hbar\right).$$

I.e. plane waves. But these are not normalisable under the $\langle\cdot|\cdot\rangle_2$ inner product:

$$\int_{-\infty}^{\infty} \left|\exp\left(ipx/\hbar\right)\right|^2 \mathrm{d}x = \infty$$

hence they do not live in a Hilbert space.

Clearly, though, they are objects of physical interest in quantum mechanics. Furthermore, they can even be made mathematically rigorous! To see this, note that, on any interval, we can actually decompose any function containing a finite number of discontinuities into the basis of plane waves, using the Fourier transform. So plane waves are mathematically and physically well defined, and it seems odd to exclude them by fiat[1].

18.2 Infinite Well

The second simplest case in QM is the infinite potential well, with

$$V(x) = \begin{cases} 0, & 0 \le x \le L \\ \infty, & \text{otherwise} \end{cases}$$

and

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$$

with

$$\langle x|\phi_n\rangle = \frac{1}{\sqrt{L}} \sin\left(\frac{n\pi x}{L}\right) = \langle x|s_n\rangle.$$

[1]Prof. Robbins comments: "I would say that $|x\rangle$ and $|p\rangle$ are well defined mathematically - they don't need to be banished from mathematically rigorous discourse. But they are not elements of L^2.

Are they well defined physically? I guess I would say they are unphysical idealizations/limits of physically meaningful states. Why unphysical? Well, for one thing, for most Hamiltonians, $|x\rangle$ and $|p\rangle$ would have infinite energy.

In the same vein, it might be worthwhile pointing out that there is a mathematical distinction between spectrum and eigenvalue. Self-adjoint operators have real spectra, but a point in the spectrum needn't be an eigenvalue (as with the position and momentum operators). This brings us back to $|x\rangle$ and $|p\rangle$ not being in the Hilbert space. One can say that for points in the spectrum, there are approximate eigenstates (eg, functions that approximate delta functions in the case of x)."

Let us revisit the issues raised in Section 8.4.3. Since the eigenenergies are non-degenerate, $|\phi_n\rangle$ must form a complete orthonormal basis for a Hilbert space. Evidently the dimension of this space is infinity. But which infinity?

Well, $n \in \mathbb{Z}$, so $\dim \mathcal{H} = \aleph_0$: it is countable. Any state $|\psi\rangle \in \mathcal{H}$ can be decomposed into the $|s_n\rangle$ basis using

$$\mathbb{I} = \sum_{n=1}^{\aleph_0} |s_n\rangle\langle s_n|$$

and

$$\langle s_n|s_m\rangle = \delta_{nm}.$$

But on the other hand, any state can also be written in the position basis, using

$$\mathbb{I} = \int_0^L |x\rangle\langle x|\mathrm{d}x$$

and

$$\langle x|y\rangle = \delta\left(x - y\right).$$

This assigns a separate state to each point on the real interval $x \in [0, L]$. But the number of points on an interval of the real line is the definition of the cardinality of the continuum, suggesting $\dim \mathcal{H} = \mathfrak{c}$. And Cantor showed that $\mathfrak{c} > \aleph_0$.

So which is it? The resolution to both of these paradoxes is that states in quantum mechanics do not actually live in Hilbert spaces after all...

18.3 Concluding remarks

Von Neumann and Dirac proposed alternative approaches to axiomatising quantum mechanics. Von Neumann's approach rejects plane waves and delta functions as non-rigorous. In his 1932 textbook he states that Hilbert spaces are the correct objects of study for quantum mechanics. But already by 1935 he had realised that this was not the case. In the subsequent century, Von

Neumann's approach has been formalised into the theory of C^*-*algebras*. Coarsely, we might say that this is the mathematicians' approach to quantum mechanics[2].

Dirac includes plane waves and delta functions, and so rejects the notion that Hilbert spaces fully describe quantum systems. Dirac's approach was later formalised using the idea of *rigged Hilbert spaces* (which, in being rigged to work formally, cease to be Hilbert spaces). Coarsely, this is the physicists' approach to quantum mechanics.

More generally, incorporating special relativity into quantum mechanics, and allowing for particle creation and annihilation, one arrives at quantum field theory (QFT). Von Neumann's approach led to *Algebraic QFT*, while Dirac's led to *Axiomatic QFT* (Wightman's axioms).

A rigorous formal axiomatisation of quantum field theory, which simultaneously allows useful calculations, is yet to be found. It is one of the Clay Mathematics Institute's million-dollar Millennium Prize Problems (Existence & the Mass Gap). You are encouraged to give it a go.

[2] Prof. Robbins comments: "I don't know how von Neumann's views evolved, but I don't think all mathematicians have given up on the Hilbert space formulation. That is, you can find mathematical treatements of nonrelativistic QM formulated in terms of Hilbert space. As I understand, the move to C^* algebras is motivated by regarding density operators rather than wavefunctions as describing the state of a physical system. This has the advantage of removing the phase ambiguity as well as allowing for mixed states alongside pure states."