2 Hilbert Space

The realm of Quantum Mechanics is Hilbert space[3], so we'll begin by exploring the properties of these. This chapter will necessarily be almost entirely mathematical; the physics comes later.

2.1 Definition of Hilbert Space

Hilbert space is a vector space $\mathcal{H}$ over $\mathbb{C}$ that is equipped with a complete inner product. Let's take a moment to understand what this means; much of it will be familiar from IB Linear Algebra or IB Methods.

Saying that $\mathcal{H}$ is a vector space means that it is a set on which we have an operation + of *addition*, obeying

$$\begin{aligned} \text{commutativity} \quad & \psi + \phi = \phi + \psi \\ \text{associativity} \quad & \psi + (\phi + \chi) = (\psi + \phi) + \chi \\ \text{identity} \quad & \exists!\ o \in \mathcal{H} \ \text{ s.t. } \ \psi + o = \psi \end{aligned} \tag{2.1}$$

for all $\psi, \phi, \chi \in \mathcal{H}$. Furthermore, since it is a vector space over $\mathbb{C}$, we can multiply our vectors by numbers $a, b, c, \ldots \in \mathbb{C}$, called scalars. This multiplication is

$$\begin{aligned} \text{distributive over } \mathcal{H} \quad & c(\psi + \phi) = c\psi + c\phi \\ \text{distributive in } \mathbb{C} \quad & (a + b)\psi = a\psi + b\psi\ . \end{aligned} \tag{2.2}$$

In addition, $\mathcal{H}$ comes equipped with an *inner product*. This is a map $(\ ,\ ) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ that obeys

$$\begin{aligned} \text{conjugate symmetry} \quad & (\phi, \psi) = \overline{(\psi, \phi)} \\ \text{linearity} \quad & (\phi, a\psi) = a\,(\phi, \psi) \\ \text{additivity} \quad & (\phi, \psi + \chi) = (\phi, \psi) + (\phi, \chi) \\ \text{positive–definiteness} \quad & (\psi, \psi) \geq 0 \ \ \forall\ \psi \in \mathcal{H}, \ \text{with equality iff } \psi = o\,. \end{aligned} \tag{2.3}$$

Note that the first two of these imply $(a\phi, \psi) = \overline{a}\,(\phi, \psi)$ so that $(\ ,\ )$ is *antilinear* in its first (leftmost) argument[4]. Note also that $(\psi, \psi) = \overline{(\psi, \psi)}$ so that this is necessarily real. Whenever we have inner product, we can define[5] the *norm* of a state to be

$$\|\psi\| = \sqrt{(\psi, \psi)}\,. \tag{2.4}$$

The properties (2.3) also ensure that the Cauchy–Schwarz inequality $|(\phi, \psi)|^2 \leq (\phi, \phi)\,(\psi, \psi)$ holds.

---

[3] In this course, we'll focus on Dirac's formulation of QM, which is based on Hilbert space. This is by far the most commonly used approach. However, there are some approaches to QM (notably deformation quantization and the theory of C^*-algebras) in which Hilbert spaces do not play a prominent role. We won't discuss any of them.

[4] In the maths literature, the inner product is often taken to be linear in the left entry and antilinear in the right. We'll follow the QM literature, which always uses the opposite convention.

[5] A norm is a more fundamental notion than an inner product: there are normed vector spaces that do not have inner products, while every inner product space has a norm, defined as in (2.4).

As always, a set of vectors $\{\phi_1, \phi_2, \ldots, \phi_n\}$ are *linearly independent* iff the only solution to

$$c_1\phi_1 + c_2\phi_2 + \cdots + c_n\phi_n = o \tag{2.5}$$

for $c_i \in \mathbb{C}$ is $c_1 = c_2 = \cdots = c_n = 0$. The *dimension* of the vector space is the largest possible number of linearly independent vectors we can find. If there is no such largest number, then we say the vector space has infinite dimension. A set of vectors $\{\phi_1, \phi_2, \ldots, \phi_n\}$ is *orthonormal* with respect to (,) if

$$(\phi_b, \phi_a) = \begin{cases} 0 & \text{when } b \neq a, \\ 1 & \text{when } b = a. \end{cases} \tag{2.6}$$

An orthonormal set $\{\phi_1, \phi_2, \ldots, \phi_n\}$ forms a *basis* of an n-dimensional Hilbert space if every $\psi \in \mathcal{H}$ can be uniquely expressed as a sum $\psi = \sum_{a=1}^{n} c_a\phi_a$, with some coefficients $c_a \in \mathbb{C}$. We can determine these coefficients by taking the inner product with ϕ_b, since

$$(\phi_b, \psi) = \left(\phi_b, \sum_a c_a\phi_a\right) = \sum_a c_a(\phi_b, \phi_a) = c_b \tag{2.7}$$

using the linearity of the inner product and orthonormality of the ϕ_a.

Quantum mechanics makes use of both finite and infinite dimensional Hilbert spaces, as we'll see. In the infinite dimensional case, we have to decide what we mean by an 'infinite linear combination' of (*e.g.* basis) vectors. Not every such infinite sum makes sense, because infinite sums such as $\sum_{a=1}^{\infty} c_a\phi_a$ might not converge. Technically, we consider the partial sums $S_N = \sum_{a=1}^{N} c_a\phi_a$ of just the first N terms, and say that the Cauchy sequence $\{S_1, S_2, \ldots\}$ *converges in the norm* if there is some vector $\Psi \in \mathcal{H}$ to which it converges, in the sense that

$$\lim_{N\to\infty} \|S_N - \Psi\| = 0\,. \tag{2.8}$$

To say that $\mathcal{H}$ is *complete* (or complete in the norm (2.4)) means that every Cauchy sequence $\{S_1, S_2, \ldots\}$ indeed converges in $\mathcal{H}$. This captures the heuristic idea that there are no points 'missing' from $\mathcal{H}$. Importantly, it also allows us to analysis: being able to differentiate vectors in $\mathcal{H}$ requires that we can take limits, and to do this we need to know whether the limits exist.

The interplay between the vector space structure and this requirement of completeness can be very subtle in infinite dimensions — you'll see much more of this if you take the Part II course on Linear Analysis or (next term) Functional Analysis. In this course, we'll largely ignore such subtleties, not because they're not interesting, but because they're a distraction from all the interesting physics we need to learn!

2.1.1 Examples

Let's now look at a few examples of Hilbert spaces, pointing out where they're relevant to physics.

The simplest case is when $\mathcal{H}$ is finite dimensional. In this case, as a vector space we have $\mathcal{H} \cong \mathbb{C}^n$ for some n. You might think we still have lots of choice in picking an inner product, but it turns out that finite dimensional Hilbert space is always isomorphic to one with inner product

$$(\mathbf{v},\mathbf{u}) = \sum_{i=1}^{n} \overline{v}_i u_i \tag{2.9}$$

and corresponding norm

$$\|\mathbf{u}\| = \sqrt{\sum_{i=1}^{n} |u_i|^2}\,. \tag{2.10}$$

Here $\{u_i\}$ are just complex numbers – the components of the vector $\mathbf{u}$ in the canonical basis.

In physics, finite dimensional Hilbert spaces often arise as idealizations or 'toy models'. We may wish to illustrate some quantum phenomenon by first considering an especially simple case that has only finitely many (perhaps just two!) different things it can do. Alternatively, we may be concerned with just a finite dimensional subspace of a larger physical system. For example, we may be interested in states at a particular energy level of a degenerate Hamiltonian, which can easily be finite dimensional. Also, we'll see later that the angular behaviour of systems with a fixed (& finite) total angular momentum is also captured by a finite dimensional Hilbert space.

A simple infinite dimensional generalisation of this is the space of infinite sequences of complex numbers $\mathbf{u} = (u_1, u_2, u_3, \ldots)$ such that

$$\sum_{i=1}^{\infty} |u_i|^2 < \infty\,. \tag{2.11}$$

This space is known as ℓ^2 and, heuristically, you can think of it as '$\mathbb{C}^n$ with $n = \infty$'. The inner product between two such sequences $\mathbf{u}$ and $\mathbf{v}$ is defined as

$$(\mathbf{v},\mathbf{u}) = \sum_{i=1}^{\infty} \overline{v_i} u_i \tag{2.12}$$

as an obvious generalisation of the finite dimensional case (2.9). The Cauchy–Schwartz inequality gives $|(\mathbf{v},\mathbf{u})| \leq \|\mathbf{v}\|\,\|\mathbf{u}\| < \infty$, so this inner product converges provided the norms do. Again, the notion of completeness of ℓ^2 with respect to this norm enables us to meaningfully take limits and, ultimately, differentiate vectors $\mathbf{u} \in \ell^2$.

In physics, ℓ^2 arises in many places. For example, the space of energy eigenstates in an infinite square well, or in the harmonic oscillator potential, may be thought of as ℓ^2. Similarly, the space of *bound* states ($E < 0$) of the hydrogen atom is ℓ^2, though the scattering states ($E \geq 0$) are not.

We'll often also meet infinite dimensional vector spaces that are spaces of functions, such as the wavefunctions you dealt with throughout IB QM. Given two functions $\psi : \mathbb{R} \to \mathbb{C}$ and $\phi : \mathbb{R} \to \mathbb{C}$, we can define linear combinations $a\psi + b\phi$ for any $a, b \in \mathbb{C}$ in the obvious way

$$(a\psi + b\phi)(x) = a\psi(x) + b\phi(x) \tag{2.13}$$

where the multiplication and addition on the rhs are just those in $\mathbb{C}$, so spaces of functions are naturally infinite dimensional vector spaces (as you saw in IB Methods). To turn this space of functions into a Hilbert space, we first give it the norm

$$\|\psi\| = \sqrt{\int_{\mathbb{R}} |\psi(x)|^2 \, \mathrm{d}x} \tag{2.14}$$

and require that $\|\psi\| < \infty$ – *i.e.* that the integral converges. In physics, functions for which $\|\psi\| < \infty$ holds are called *normalisable.* Note that just asking for (2.14) to converge is not a very strong restriction, and so our normed function space contains a very wide class of functions, including all piecewise continuous functions that decay sufficiently rapidly as $|x| \to \infty$, and even some functions that are singular for some discrete values of x, provided these singularities are not strong enough to cause the *integral* (2.14) to diverge[6].

Finally, we take the inner product between two functions ψ, ϕ whose norms are finite to be

$$(\phi, \psi) = \int_{\mathbb{R}} \bar{\phi}(x)\psi(x) \, \mathrm{d}x \tag{2.15}$$

and again this converges by Cauchy–Schwarz. In physics, we often call $\psi(x)$ the *wavefunction* of the particle, and sometimes call (ϕ, ψ) the *overlap integral* between two wavefunctions.

2.1.2 Dual Spaces

As with any vector space, the *dual* $\mathcal{H}^*$ of a Hilbert space $\mathcal{H}$ is the space of linear maps $\mathcal{H} \to \mathbb{C}$. That is, an element $\varphi \in \mathcal{H}^*$ defines a map $\varphi : \psi \mapsto \varphi(\psi) \in \mathbb{C}$ for every $\psi \in \mathcal{H}$, such that

$$\varphi : a\psi_1 + b\psi_2 \mapsto a\varphi(\psi_1) + b\varphi(\psi_2) \tag{2.16}$$

for all $\psi_1, \psi_2 \in \mathcal{H}$ and $a, b \in \mathbb{C}$.

One way to construct such a map is to use the inner product: given some state $\phi \in \mathcal{H}$ we can define an element $(\phi, \quad) \in \mathcal{H}^*$ which acts by

$$(\phi, \quad) : \psi \mapsto (\phi, \psi) \tag{2.17}$$

i.e. we take the inner product of $\psi \in \mathcal{H}$ with our chosen element ϕ. The linearity properties of the inner product transfer to ensure that $(\phi, \quad)$ is indeed a *linear* map — note that since the inner product is antilinear in its first entry, it's important that our chosen element ϕ

---

[6] The requirement that the space be *complete* in the norm (2.14) is rather subtle. If $\|\psi - \phi\| = 0$, then we must identify ψ and ϕ as *the same object* in our space. This does not necessarily mean that they're identical as functions, because *e.g.* they could take different at some discrete points $x_i \subset \mathbb{R}$, as the non-zero value of $\psi - \phi$ at these discrete points would not contribute to (2.14). In particular, any function that is non-zero only at a discrete set of points should be identified with the zero function. The resulting space is known as $L^2(\mathbb{R}, \mathrm{d}x)$ or sometimes just L^2 for short. (The L stands for *Lebesgue*, and is an example of a more general type of normed function space.) $L^2(\mathbb{R}, \mathrm{d}x)$ consists of *equivalence classes* of Cauchy sequences of functions that are convergent in the norm (2.14). In this course we'll mostly gloss over such technicalities, and they're certainly non-examinable. For a deeper discussion of Hilbert space, see the Part II Linear Analysis & Functional Analysis courses.

sits in this first entry. In 1B Linear Algebra you proved that, in finite dimensions, *every* element of $\mathcal{H}^*$ arises this way. That is, *any* linear map $\varphi : \mathcal{H} \to \mathbb{C}$ can be written as $(\phi, \;\;)$ for some fixed choice of $\phi \in \mathcal{H}$. This means in particular that the inner product $(\;,\;)$ provides a vector space isomorphism

$$\mathcal{H}^* \overset{(\,,\,)}{\cong} \mathcal{H}\,. \tag{2.18}$$

Comfortingly, the same result also holds in infinite dimensions[7], but it's non–trivial to prove and is known as the *Riesz Representation Theorem*. The isomorphism $\mathcal{H}^* \cong \mathcal{H}$ is what is special about Hilbert spaces among various other infinite dimensional vector spaces, and makes them especially easy to handle.

2.1.3 Dirac Notation and Continuum States

From now on, in this course we'll use a notation for Hilbert spaces that was introduced by Dirac and is standard throughout the theoretical physics literature. Dirac denotes an element of $\mathcal{H}$ as $|\psi\rangle$, where the symbol "$|\;\;\rangle$" is known as a *ket*. An element of the dual space is written $\langle\phi|$ and the symbol "$\langle\;\;|$" is called a *bra*. The relation between the ket $|\phi\rangle \in \mathcal{H}$ and the bra $\langle\phi| \in \mathcal{H}^*$ is what we would previously have written as ϕ *vs* $(\phi, \;\;)$. The inner product between two states $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ is then written $\langle\phi|\psi\rangle$ forming a *bra-ket* or bracket. Note that this implicitly uses the isomorphism $\mathcal{H}^* \cong \mathcal{H}$ provided by the inner product, building it into the notation. Recall also that (at least in all physics courses!) $\langle\phi|\psi\rangle$ is *anti*linear in $|\phi\rangle$.

Given an orthonormal basis $\{|e_a\rangle\}$ of $\mathcal{H}$, at least for the cases $\mathcal{H} \cong \mathbb{C}^n$ or $\mathcal{H} \cong \ell^2$, in Dirac notation we can expand a general ket $|\psi\rangle$ as

$$|\psi\rangle = \sum_a \psi_a |e_a\rangle \tag{2.19}$$

in terms of this basis. Then $\langle\chi|\psi\rangle = \sum_{a,b} \overline{\chi}_b \psi_a \langle e_b|e_a\rangle = \sum_a \overline{\chi}_a \psi_a$ as usual.

It's very useful to be able to extend this idea also to function spaces. In this case, we introduce a 'continuum basis' with elements $|a\rangle$ labelled by a continuous variable a, normalised so that

$$\langle a'|a\rangle = \delta(a' - a) \tag{2.20}$$

using the Dirac δ-function. Then we write

$$|\psi\rangle = \int \psi(a)|a\rangle\, da \tag{2.21}$$

---

[7] We should really be more careful here, though we won't be concerned with the following subtleties in this course. In infinite dimensions we distinguish the *algebraic dual* space – the space of *all* linear functionals $\varphi : \mathcal{H} \to \mathbb{C}$ from the *continuous dual*, where the number $\varphi(\psi)$ is required to be vary continuously as ψ varies in $\mathcal{H}$. The Riesz Representation applies to the continuous dual. For example, the algebraic dual also contains distributions such as the Dirac δ, acting as $\delta : \psi \mapsto \delta[\psi] \equiv \psi(0)$ " $= \int_{\mathbb{R}} \delta(x)\, \psi(x)\, dx$", with $\psi \in L^2(\mathbb{R}, dx)$. Certainly $\delta(x)$ is not itself square-integrable, so is not in the Hilbert space. However, since $L^2(\mathbb{R}, dx)$ contains discontinuous functions, nor does $\delta[\psi]$ vary smoothly with ψ. Functional analysis is almost always interested in the continuous dual, and this if often called just the *dual* space.

to expand a general $|\psi\rangle$ in terms of the $|a\rangle$'s. The point of the normalization (2.21) is that

$$\langle\chi|\psi\rangle = \int \overline{\chi(b)}\psi(a)\,\langle b|a\rangle\,db\,da = \int \overline{\chi(b)}\psi(a)\,\delta(b-a)\,db\,da = \int \overline{\chi(a)}\psi(a)\,da \tag{2.22}$$

which is just the inner product (and also norm) we gave $L^2(\mathbb{R}, da)$ before. Indeed, a key example of a 'continuum basis' is the *position* basis $\{|x\rangle\}$, where $x \in \mathbb{R}$. Expanding a general state $|\psi\rangle$ as an integral

$$|\psi\rangle = \int_{\mathbb{R}} \psi(x')|x'\rangle\,dx'\,, \tag{2.23}$$

we see that the complex coefficients are

$$\langle x|\psi\rangle = \int_{\mathbb{R}} \psi(x')\,\langle x|x'\rangle\,dx' = \psi(x). \tag{2.24}$$

In other words, the position space wavefunctions we're familiar with are nothing but the coefficients of a state $|\psi\rangle \in \mathcal{H}$ in a particular (position) continuum basis.

As always, we could equally choose to expand this same vector in any number of different bases. For example, our state $|\psi\rangle = \int_{\mathbb{R}} \psi(x)|x\rangle\,dx$ from above can equally be expanded in the momentum basis as $|\psi\rangle = \int_{\tilde{\mathbb{R}}} \tilde{\psi}(p)|p\rangle\,dp$, where the new coefficients $\tilde{\psi}(p) = \langle p|\psi\rangle$ are the *momentum space wavefunction*. Later, we'll show that $\langle x|p\rangle = \mathrm{e}^{\mathrm{i}xp/\hbar}/\sqrt{2\pi\hbar}$, so these two sets of coefficients are related by

$$\begin{aligned}\langle x|\psi\rangle &= \int_{\tilde{\mathbb{R}}} \tilde{\psi}(p)\,\langle x|p\rangle\,dp = \frac{1}{\sqrt{2\pi\hbar}}\int_{\tilde{\mathbb{R}}} \mathrm{e}^{\mathrm{i}xp/\hbar}\,\tilde{\psi}(p)\,dp \\ \langle p|\psi\rangle &= \int_{\mathbb{R}} \psi(x)\,\langle p|x\rangle\,dx = \frac{1}{\sqrt{2\pi\hbar}}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}xp/\hbar}\,\psi(x)\,dx\,.\end{aligned} \tag{2.25}$$

(In the first line here, we expanded $|\psi\rangle$ in the momentum basis, then took the inner product with $|x\rangle$, while in the second we expanded the same $|\psi\rangle$ in the postion basis and took the inner product with $|p\rangle$.) Equation (2.25) is just the statement that the position and momentum space wavefunctions are eachother's Fourier transforms, again familiar from IB QM.

The real point I wish to make is that the fundamental object is the abstract vector $|\psi\rangle \in \mathcal{H}$. All the physical information about a quantum system is encoded in its state vector $|\psi\rangle$; the wavefunctions $\psi(x)$ or $\tilde{\psi}(p)$ are merely the expansion coefficients in some basis. Like any other choice of basis, this expansion may be useful for some purposes and unhelpful for others.

Finally, a technical[8] point. Although using 'continuum bases' such as $\{|x\rangle\}$ or $\{|p\rangle\}$ is convenient, because it emphasizes the similarities between the infinite and finite–dimensional cases, it blurs the analysis. In particular, if $\langle x'|x\rangle = \delta(x'-x)$ then the norm

$$\||x\rangle\|^2 = \delta(x-x) = \delta(0) \qquad ??? \tag{2.26}$$

---

[8] And definitely non-examinable, at least in this course.

so, whatever these objects $|x\rangle$ are, they certainly do not lie in our Hilbert space. It is possible to make good mathematical sense of these by appropriately enlarging our spaces to include spaces of distributions, but for the most part (and certainly in this course) physicists are content to say that continuum states such as $|x\rangle$ are allowable as basis elements, but call them *non–normalizable states*: actual physical particles are never represented by a non-normalizable state. If you're interested in a more detailed discussion, I recommend the book by Hall listed above.

2.2 Operators

A *linear operator* A is a map[9] $A : \mathcal{H} \to \mathcal{H}$ that is compatible with the vector space structure in the sense that

$$A(c_1|\phi_1\rangle + c_2|\phi_2\rangle) = c_1 A|\phi_1\rangle + c_2 A|\phi_2\rangle\,, \tag{2.27}$$

in Dirac notation. All the operators we meet in Quantum Mechanics will be linear, so henceforth we'll just call them 'operators'. Operators form an *algebra*: Given two such linear operators A, B, we define their sum $\alpha A + \beta B$ as

$$(\alpha A + \beta B) : |\phi\rangle \mapsto \alpha A|\phi\rangle + \beta B|\phi\rangle \tag{2.28}$$

for all $\alpha, \beta \in \mathbb{C}$ and all $|\phi\rangle \in \mathcal{H}$, and take their product AB to be the composition

$$AB : \phi \mapsto A \circ B|\phi\rangle = A(B|\phi\rangle) \tag{2.29}$$

for all $|\phi\rangle \in \mathcal{H}$. One can check that both the sum and product of two linear operators is again a linear operator in the sense of (2.27)[10].

The operator algebra is associative, so $A(BC) = (AB)C$, but not commutative and in general $AB \neq BA$ so that the order in which the operators act is important. The difference between these two actions is known as the *commutator*

$$[A, B] = AB - BA\,. \tag{2.30}$$

---

[9] Here we are being a little sloppy. In the infinite dimensional case, it often happens that operators are not defined on the whole of $\mathcal{H}$, but just on some domain $D(A) \subset \mathcal{H}$ which depends on the operator itself. For example, the momentum operator P acts on position space wavefunctions $\psi(x) \in L^2(\mathbb{R}, \mathrm{d}x)$ by $-\mathrm{i}\hbar\,\partial/\partial x$. If $\psi(x)$ is discontinuous, then it could be that $\int |\psi'|^2\,\mathrm{d}x$ diverges even though $\int |\psi|^2\,\mathrm{d}x$ itself is finite. Understanding the correct domain of various operators is an important part of functional analysis. We'll largely ignore such subtleties in this course.

[10] A linear operator is *bounded* if, for all $|\psi\rangle \in \mathcal{H}$,

$$\|A|\psi\rangle\| \leq M\,\||\psi\rangle\| \qquad \text{for some fixed } M > 0\,.$$

The space of such bounded linear operators on $\mathcal{H}$ is denoted as $\mathcal{B}(\mathcal{H})$. Bounded linear operators thus map normalisable states to normalisable states, and so act on the whole Hilbert space $\mathcal{H}$. This usually makes them the nicest ones to deal with. One of the reasons we'll be largely ignoring the analysis aspects of Hilbert space in this course is that the operators we commonly deal with in quantum mechanics, such as the position and momentum operators $\mathbf{X}$ and $\mathbf{P}$, are *unbounded*. (For example, even when its wavefunction is correctly normalised, a particle can be located at arbitrarily large $\mathbf{x} \in \mathbb{R}^3$, or have arbitrarily large momentum.) It's of course possible to handle such unbounded operators rigorously, but doing so involves an extra layer of technicality that would take us too far afield here. For further discussion, see that Part II course on Analysis of Functions, or the recommended book by Hall.

This commutator obeys the following properties:

$$\begin{aligned} \text{antisymmetry} \quad & [A,B] = -[B,A] \\ \text{linearity} \quad & [\alpha_1 A_1 + \alpha_2 A_2, B] = \alpha_1[A_1,B] + \alpha_2[A_2,B] \qquad \forall\ \alpha_1,\alpha_2 \in \mathbb{C} \\ \text{Leibniz identity} \quad & [A,BC] = [A,B]C + B[A,C] \\ \text{Jacobi identity} \quad & [A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0\,. \end{aligned} \tag{2.31}$$

Each of these may easily be verified from the definitions (2.30) & (2.27). Commutators of operators arise often in quantum mechanics, as I'm sure you remember from 1B QM.

A state $\psi \in \mathcal{H}$ is said to be an *eigenstate* of an operator A if

$$A|\psi\rangle = a_\psi|\psi\rangle \tag{2.32}$$

where the number $a_\psi \in \mathbb{C}$ is known as the *eigenvalue*. Thus, if $|\psi\rangle$ is an eigenstate of A, acting on $|\psi\rangle$ with A returns exactly the same vector, except that it is rescaled by a number that may depend both on the state and the particular operator. The set of all eigenvalues of an operator A is sometimes called the *spectrum* of A, while the number of linearly independent eigenstates having the same eigenvalue is called the *degeneracy* of this eigenvalue.

One place where Dirac notation is particularly convenient is that it allows us to label states by their eigenvalues. We let $|q\rangle$ denote an eigenstate of some operator Q with eigenvalue q, so that $Q|q\rangle = q|q\rangle$. For example, a state that is an eigenstate of the position operator $\mathbf{X}$ with position $\mathbf{x} \in \mathbb{R}^3$ (representing a particle that is definitely located at $\mathbf{x}$) is written $|\mathbf{x}\rangle$, while the state $|\mathbf{p}\rangle$ is an eigenstate of the momentum operator $\mathbf{P}$ with eigenvalue $\mathbf{p}$. There's a potential confusion here: $|\mathbf{x}\rangle$ does *not* refer to the function $\mathbf{x}$, but rather a state in $\mathcal{H}$ whose only support is at $\mathbf{x}$. We saw above that the wavefunction corresponding to $|\mathbf{x}\rangle$ was really a δ-function.

In this course, I'll mostly use the (unconventional!) convention that operators are written using capital letters, with their eigenvalues being labelled by the same letter in lowercase. However, I won't stick to this religiously; a notable and deeply ingrained exception is to use E for the eigenvalues of the Hamiltonian operator H.

The extra structure of the inner product on $\mathcal{H}$ allows us to also define the *adjoint* $A^\dagger$ of an operator A by

$$\langle\phi|A^\dagger|\psi\rangle = \overline{\langle\psi|A|\phi\rangle} \qquad \text{for all } |\phi\rangle, |\psi\rangle \in \mathcal{H}\,. \tag{2.33}$$

One can easily[11] check that

$$(A+B)^\dagger = A^\dagger + B^\dagger\,, \qquad (AB)^\dagger = B^\dagger A^\dagger\,, \qquad (\alpha A)^\dagger = \overline{\alpha} A^\dagger \qquad \text{and} \qquad (A^\dagger)^\dagger = A\,, \tag{2.34}$$

from which it follows that $[A,B]^\dagger = [B^\dagger, A^\dagger]$. Note also that the adjoint of the eigenvalue equation $A|a\rangle = a|a\rangle$ is

$$\langle a|A^\dagger = \langle a|\overline{a}\,, \tag{2.35}$$

---

[11] At least in the finite dimensional case. We'll assume without being careful that it holds in the infinite dimensional case too.

where $A^\dagger$ is taken to act to the left on the dual vector $\langle a|$.

An operator Q is called *Hermitian*[12] if $Q^\dagger = Q$, so that $\langle\phi|Q|\psi\rangle = \overline{\langle\psi|Q|\phi\rangle}$ for all $|\phi\rangle, |\psi\rangle \in \mathcal{H}$. Hermitian operators are very special and have a number of important properties. Firstly, suppose $|q\rangle$ is an eigenstate of a Hermitian operator Q, with eigenvalue q. Then

$$q\langle q|q\rangle = \langle q|Q|q\rangle = \overline{\langle q|Q|q\rangle} = \overline{q}\langle q|q\rangle\,, \tag{2.36}$$

so $\overline{q} = q$ and the eigenvalues of a Hermitian operator are real. Second, suppose $|q_1\rangle$ and $|q_2\rangle$ are both eigenstates of Q with distinct eigenvalues q_1 and q_2. Then

$$(q_1 - q_2)\langle q_1|q_2\rangle = \langle q_1|Q^\dagger|q_2\rangle - \langle q_1|Q|q_2\rangle = 0\,, \tag{2.37}$$

where the first equality uses the reality of q_1 and the second uses the fact that Q is Hermitian. Since $q_1 \neq q_2$ by assumption, we must have $\langle q_1|q_2\rangle = 0$, so eigenstates of Hermitian operators with distinct eigenvalues are orthogonal and, perhaps after rescaling, we can choose them to be orthonormal.

In the finite dimensional case, you proved in IB Linear Algebra that the set of eigenvectors of a given Hermitian operator form a basis of $\mathcal{H}$; that is, they form a complete, orthonormal set. This property allows us to express Hermitian operators in a form that is often useful. If $\{|n\rangle\}$ is an orthonormal basis of eigenstates of a Hermitian operator Q, with eigenvalues $\{q_n\}$, then we can write

$$Q = \sum_n q_n\,|n\rangle\langle n| \tag{2.38}$$

where we think of this as acting on some $|\psi\rangle \in \mathcal{H}$ by

$$Q|\psi\rangle = \sum_n q_n\,|n\rangle\langle n|\psi\rangle\,. \tag{2.39}$$

Note that indeed $Q|\psi\rangle \in \mathcal{H}$, since the $\langle n|\psi\rangle$ are just complex numbers, whereas each term in the sum also involves $|n\rangle \in \mathcal{H}$. In particular, expressing $|\psi\rangle$ in this basis as $|\psi\rangle = \sum_m c_m\,|m\rangle$ gives

$$Q|\psi\rangle = \sum_{n,m} q_n c_m\,|n\rangle\,\langle n|m\rangle = \sum_n q_n c_n\,|n\rangle \tag{2.40}$$

---

[12] The terminology 'Hermitian' is used in the physics literature, coming from the fact that if $\mathcal{H}$ is finite dimensional, we can represent linear operators obeying $Q^\dagger = Q$ by Hermitian matrices. In functional analysis, we'd have to be a little more careful about exactly which states we allow our operator to act on. For example, differentiating a non-smooth but square–integrable function will typically sharpen its singularities and may render it non-normalisable. Thus the momentum operator $-\mathrm{i}\hbar d/dx$ may take a state in $\mathcal{H} \cong L^2(\mathbb{R}, dx)$ out of $L^2(\mathbb{R}, dx)$, and to keep analytic control we should first limit the states on which we allow the momentum operator to act. Being more careful, we'd say that operators obeying $\langle\psi|Q|\phi\rangle = \overline{\langle\phi|Q|\psi\rangle}$ for all $|\phi\rangle$, $|\psi\rangle \in \mathcal{H}$ are *symmetric*, while symmetric operators for which the domain $\mathrm{Dom}(Q) \subseteq \mathcal{H}$ of the operator is the same as the domain $\mathrm{Dom}(Q^\dagger)$ of the adjoint are called *self-adjoint*. The distinction between operators that are self-adjoint and those that are merely symmetric is a little technical, but has important consequences for their spectrum and the existence of eigenstates. See *e.g.* this Wikipedia page for a basic discussion with examples. We'll largely ignore such subtleties in our course.

assuming the basis is orthonormal.

One reason this representation of Q is useful is that it allows us to define functions of operators. We set

$$f(Q) = \sum_n f(q_n)\,|n\rangle\langle n| \tag{2.41}$$

which makes sense provided $f(q_n)$ is defined for all eigenvalues q_n of the original operator. For example, the inverse of an operator Q is the operator

$$Q^{-1} = \sum_n \frac{1}{q_n}\,|n\rangle\langle n| \tag{2.42}$$

which makes sense provided $q_n \neq 0$ for all n, while

$$\ln Q = \sum_n \ln(q_n)\,|n\rangle\langle n| \tag{2.43}$$

is meaningful whenever Q has eigenvalues that are all positive–definite.

A particularly important example of this is that we can represent the identity operator $1_{\mathcal{H}}$ on $\mathcal{H}$ as

$$1_{\mathcal{H}} = \sum_n |n\rangle\langle n| \qquad \text{or} \qquad 1_{\mathcal{H}} = \int |q\rangle\langle q|\,dq \tag{2.44}$$

where $\{|n\rangle\}$ or $\{|q\rangle\}$ are any bases of $\mathcal{H}$, respectively for discrete or continuous labels. This is often convenient for moving between different bases. For example, if $\{|p\rangle\}$ denotes a complete set of eigenstates of the momentum operator P, with $P|p\rangle = p|p\rangle$, then to represent this operator acting on a state $|\psi\rangle$ in the position basis we write

$$\begin{aligned}\langle x|P|\psi\rangle &= \int \langle x|P|p\rangle\langle p|\psi\rangle\,dp = \int \langle x|p\rangle\,p\,\tilde{\psi}(p)\,dp \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int \mathrm{e}^{ixp/\hbar}\,p\,\tilde{\psi}(p)\,dp = -\mathrm{i}\hbar\frac{\partial}{\partial x}\left[\frac{1}{\sqrt{2\pi\hbar}}\int \mathrm{e}^{ixp/\hbar}\,\tilde{\psi}(p)\,dp\right]\,.\end{aligned} \tag{2.45}$$

In the first equality here we've squeezed the identity operator $1_{\mathcal{H}} = \int |p\rangle\langle p|\,dp$ in between the momentum operator P and the general state $|\psi\rangle$, then used the fact that $|p\rangle$ is an eigenstate of P, and finally used our previous expression for $\langle x|p\rangle$ and standard properties of Fourier transforms. We recognise the integral in the final expression as the position space wavefunction of $|\psi\rangle$, so altogether we have

$$\langle x|P|\psi\rangle = -\mathrm{i}\hbar\frac{\partial}{\partial x}\langle x|\psi\rangle\,, \tag{2.46}$$

or $\hat{P}\psi(x) = -\mathrm{i}\hbar\partial\psi/\partial x$ in 1B notation. This and similar manipulations will be used repeatedly throughout the course, so it's important to make sure you're comfortable with them.

If $\mathcal{H}$ is finite dimensional, we can represent any linear operator on $\mathcal{H}$ by a matrix. We pick an orthonormal basis $\{|n\rangle\}$ and define

$$A_{nm} := \langle n|A|m\rangle \tag{2.47}$$

as usual. In particular, inserting the identity operator in the form $I = \sum_n |n\rangle\langle n|$ shows that the matrix elements of the product operator AB are

$$
\begin{aligned}
(AB)_{km} = \langle k|AB|m\rangle &= \langle k|A\left(\sum_n |n\rangle\langle n|\right) B|m\rangle \\
&= \sum_n \langle k|A|n\rangle\langle n|B|m\rangle = \sum_n A_{kn}B_{nm} \,,
\end{aligned}
\tag{2.48}
$$

which just corresponds to usual matrix multiplication. Hermitian operators are represented by a Hermitian matrices, *i.e.* ones whose elements obey $(A^\dagger)_{nm} = \overline{A_{mn}}$, hence the name[13].

2.3 Composite Systems

We often have to deal with systems with more than a single degree of freedom. This could be just a single particle moving in three dimensions rather than one, or a composite system such as a Hydrogen atom consisting of both an electron and a proton, or a diamond which consists of a very large number of carbon atoms. In the case of a macroscopic object, our system could be made up of a huge number of constituent parts, each able to behave differently, at least in principle.

It should be intuitively clear that the more complicated our system is, meaning the more independent degrees of freedom it possesses, the larger a Hilbert space we'll need if we wish to encode all its possible states. We'll now understand how quantum mechanics handles systems with more than one degree of freedom by taking the tensor product of the Hilbert spaces of its constituent parts.

2.3.1 Tensor Product of Hilbert Spaces

Hilbert spaces are vector spaces over $\mathbb{C}$, so we can try to define their tensor product in the same way as we would for any pair of vector spaces. Recall from IB Linear Algebra that if $\mathcal{H}_1$ and $\mathcal{H}_2$ are two, finite dimensional Hilbert spaces, with $\{|e_a\rangle\}_{a=1}^m$ a basis of $\mathcal{H}_1$ and $\{|f_\alpha\rangle\}_{\alpha=1}^n$ a basis of $\mathcal{H}_2$, then the *tensor product* $\mathcal{H}_1 \otimes \mathcal{H}_2$ is again a vector space over $\mathbb{C}$ spanned by all pairs of elements $|e_a\rangle \otimes |f_\alpha\rangle$ chosen from the two bases[14]. It follows that

$$
\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim(\mathcal{H}_1)\,\dim(\mathcal{H}_2)
\tag{2.49}
$$

---

[13] Though see the previous footnote for subtleties in the infinite dimensional case.

[14] If you're a purist (and the PQM Tripos examiners are *not*), then there's something slightly unsatisfactory about this definition of the tensor product vector space, which is that it apparently depends on a choice of basis on both $\mathcal{H}_1$ and $\mathcal{H}_2$. We can do better, just using the abstract vector space structure, as follows. For any pair of vector spaces V and W, we first define the *free vector space* $F(V \times W)$ to be the set of all possible pairs of elements chosen from V and W. This is a vector space with the sum of a pair (v, w) and a pair (v', w') denoted simply by $(v, w) + (v', w')$ (Do not confuse the pair $(v, w) \in F(V \times W)$ with an inner product! Here v and w live in different spaces.) By itself, $F(V \times W)$ does not care that V and W are already vector spaces. For example, if $\{e_i\}$ is a basis for V and $v = c_1e_1 + c_2e_2$ for some coefficients $c_{1,2}$, the elements (v, w) and $c_1(e_1, w) + c_2(e_2, w)$ are nonetheless distinct in $F(V \times W)$. We'd like to teach the free vector space about the structure inherited from V and W individually. We thus define the tensor product $V \otimes W$ to be the *quotient* $F(V \times W)/\sim$ under the equivalence relations

$$
(v, w) + (v', w) \sim (v + v', w)\,, \qquad (v, w) + (v, w') \sim (v, w + w') \qquad \text{and} \qquad c(v, w) \sim (cv, w) \sim (v, cw)
$$

for all $v \in V$, $w \in W$ and all $c \in \mathbb{C}$. We denote the equivalence class of the pair (v, w) by $v \otimes w$.

provided these are finite. This should be familiar from IB Linear Algebra.

It's important to stress that, as with any vector space, a general element of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a *linear combination* of these basis elements. In particular, although general elements of $\mathcal{H}_1$ and $\mathcal{H}_2$ can be written

$$|\psi_1\rangle = \sum_{a=1}^{m} c_a |e_a\rangle \qquad \text{and} \qquad |\psi_2\rangle = \sum_{\alpha=1}^{m} d_\alpha |f_\alpha\rangle\,, \tag{2.50}$$

it is *not* true that a general element of $\mathcal{H}_1 \otimes \mathcal{H}_2$ necessarily takes the form $|\psi_1\rangle \otimes |\psi_2\rangle$. Rather, a general element of $\mathcal{H}_1 \otimes \mathcal{H}_2$ may be written as

$$|\Psi\rangle = \sum_{a,\alpha} r_{a\alpha}\, |e_a\rangle \otimes |f_\alpha\rangle \tag{2.51}$$

In particular, elements of the form $|\psi_1\rangle \otimes |\psi_2\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ are specified by only $\dim \mathcal{H}_1 + \dim \mathcal{H}_2$ complex coefficients, vastly fewer than is required to specify a generic element (2.51). Elements of the form $|\psi\rangle \otimes |\phi\rangle$ are sometimes called *simple*, while in physics generic elements of the form (2.51) are said to be *entangled*. In section **??** we'll explore some of the vast resources this entanglement opens up to quantum mechanics; you'll see much more of it next term if you take the course on Quantum Information & Computation.

In order to make $\mathcal{H}_1 \otimes \mathcal{H}_2$ a Hilbert space, rather than just a vector space, we must give it an inner product. We do this in the obvious way: if $\langle\ |\ \rangle_1$ and $\langle\ |\ \rangle_2$ denote inner products on $\mathcal{H}_1$ and $\mathcal{H}_2$, we first define the inner product between each pair of basis elements of $\mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$(\langle e_a| \otimes \langle f_\alpha|)\;(|e_b\rangle \otimes |f_\beta\rangle) = \langle e_a|e_b\rangle_1\, \langle f_\alpha|f_\beta\rangle_2\,, \tag{2.52}$$

and then extend it to arbitrary states $|\Psi\rangle$ by linearity. Note in particular that if $\{|e_a\rangle\}$ and $\{|f_\alpha\rangle\}$ are orthonormal bases of their respective Hilbert spaces, then $\{|e_a\rangle \otimes |f_\alpha\rangle\}$ will also be an orthonormal basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$.

The most common occurrence of this is simply a single particle moving in more than one dimension. For example, suppose a quantum particle moves in $\mathbb{R}^2$, described by Cartesian coordinates (x, y). If $\{|x\rangle\}_{x\in\mathbb{R}}$ is a complete set of position eigenstates for the x-direction, and $\{|y\rangle\}_{y\in\mathbb{R}}$ likewise a complete set for the y-direction, then the state of the particle can be expanded as

$$|\psi\rangle = \int_{\mathbb{R}\times\mathbb{R}} \psi(x,y)\, |x\rangle \otimes |y\rangle\; dx\, dy \tag{2.53}$$

where $\{|x\rangle \otimes |y\rangle\}_{x,y\,\in\mathbb{R}}$ form a continuum basis of the tensor product space. Note that in general, our wavefunction is not a simple product $\psi(x,y) \neq \psi_1(x)\psi_2(y)$ of wavefunctions in the two directions separately, though it may be possible to write it as the sum of such products, as you met repeatedly when studying separation of variables in IB Methods. In this continuum case, the inner product between any two states is

$$\begin{aligned}\langle\chi|\psi\rangle &= \int_{\mathbb{R}^2\times\mathbb{R}^2} \overline{\chi(x',y')}\,\psi(x,y)\,\langle x'|x\rangle\,\langle y'|y\rangle\;dx'\,dy'\,dx\,dy \\ &= \int_{\mathbb{R}^2\times\mathbb{R}^2} \overline{\chi(x',y')}\,\psi(x,y)\,\delta(x-x')\,\delta(y-y')\;dx'\,dy'\,dx\,dy \\ &= \int_{\mathbb{R}^2} \overline{\chi(x,y)}\,\psi(x,y)\;dx\,dy\,.\end{aligned} \tag{2.54}$$

This is exactly the structure of $L^2(\mathbb{R}^2\, d^2x)$, and we identify[15] $L^2(\mathbb{R}^2, d^2x) \cong L^2(\mathbb{R}, dx) \otimes L^2(\mathbb{R}, dy)$.

More commonly, physics considers quantum systems that live in $\mathbb{R}^3$, described by a Hilbert space $\mathcal{H} \cong L^2(\mathbb{R}^3, d^3x)$ obtained from the tensor product of three copies of the Hilbert space for a one-dimensional system. We can expand a general state in this Hilbert space as

$$|\psi\rangle = \int_{\mathbb{R}^3} \psi(x,y,z)\;|x\rangle\otimes|y\rangle\otimes|z\rangle\;dx\,dy\,dz = \int_{\mathbb{R}^3} \psi(\mathbf{x})\,|\mathbf{x}\rangle\;d^3x \tag{2.56}$$

where the last expression is just a convenient shorthand. Going further, to describe a composite system such as a Hydrogen atom that consists of two particles (an electron and proton), we need to take the tensor product of the individual Hilbert spaces of each particle. Thus (neglecting the spin of the electron and proton) the atom is described by a state

$$|\Psi\rangle = \int_{\mathbb{R}^6} \Psi(\mathbf{x}_{\rm e},\mathbf{x}_{\rm p})\;|\mathbf{x}_e\rangle\otimes|\mathbf{x}_p\rangle\;\mathrm{d}^3x_{\rm e}\,d^3x_{\rm p} \tag{2.57}$$

where $\Psi(\mathbf{x}_{\rm e},\mathbf{x}_{\rm p}) = (\langle\mathbf{x}_e|\otimes\langle\mathbf{x}_p|)|\Psi\rangle$ is the amplitude for the atom's electron to be found at $\mathbf{x}_{\rm e}$ while its proton is at $\mathbf{x}_{\rm p}$.

---

[15] There's actually one more – again non-examinable! – step in the infinite dimensional case: we must take the *completion* of $L^2(\mathbb{R}, dx) \otimes L^2(\mathbb{R}, dy)$ in the norm associated to this inner product. There is no problem if our general state $\psi(x,y)$ can be written as a *finite* sum of products of square-integrable wavefunctions, say $\phi_a(x)$ and $\rho_b(y)$ in each of the two directions separately, for then

$$\begin{aligned}\int_{\mathbb{R}^2} |\psi(x,y)|^2\,dx\,dy &= \int_{\mathbb{R}^2} \sum_{a,a',b,b'} \overline{\phi_{a'}(x)\rho_{b'}(y)}\,\phi_a(x)\rho_b(y)\,dx\,dy \\ &= \left[\sum_{a,a'}\int_{\mathbb{R}} \overline{\phi_{a'}(x)}\,\phi_a(x)\,dx\right]\left[\sum_{b,b'}\int_{\mathbb{R}} \overline{\rho_{b'}(y)}\,\rho_b(y)\,dy\right] < \infty\end{aligned}$$

with convergence guaranteed by the properties of the inner product on each copy of $L^2(\mathbb{R}, dx)$. However, there are functions in $L^2(\mathbb{R}^2, d^2x)$ that cannot be so expressed as a finite sum. A simple example is the function $f:\mathbb{R}^2\to\mathbb{C}$ given by

$$f:(x,y)\mapsto\begin{cases}1 & \text{when } \sqrt{x^2+y^2}\le a \\ 0 & \text{else}\end{cases}\,. \tag{2.55}$$

It is clear that this function is square-integrable over $\mathbb{R}^2$, but it cannot be written as a *finite* sum of products of functions of x and y separately. It thus lies in the completion of$L^2(\mathbb{R}, dx) \otimes L^2(\mathbb{R}, dy)$, but not in this space itself. The Hilbert space completion of $\mathcal{H}_1\otimes\mathcal{H}_2$ is often denoted $\mathcal{H}_1\hat{\otimes}\mathcal{H}_2$, so more correctly we have $L^2(\mathbb{R}^2, d^2x) \cong L^2(\mathbb{R}, dx)\hat{\otimes}L^2(\mathbb{R}, dy)$. We'll ignore such subtleties in this course, saving a proper treatment for the Functional Analysis course.

Once we know the Hilbert spaces appropriate to describe the fundamental constituents of our system, we can build up the Hilbert space for the combined system by taking tensor products. We should then ask *'What is the correct Hilbert space to use to describe the fundamental particles in our system?'*. Ultimately, this question can only be determined by carrying out an experiment. For example, experiments performed by Stern & Gerlach (see section 5.3.2) showed that a single electron is in fact described by a Hilbert space $\mathcal{H}_{\rm e} = L^2(\mathbb{R}^3, d^3x) \otimes \mathbb{C}^2$, formed from the tensor product of the electron's position space wavefunction with the two-dimensional Hilbert space $\mathbb{C}^2$ describing the electron's 'internal' degrees of freedom. Thus, letting $\{|\uparrow\rangle, |\downarrow\rangle\}$ be an orthonormal basis of $\mathbb{C}^2$, a generic state of the electron (whether or not it's part of a Hydrogen atom) is

$$|\Psi\rangle = |\phi\rangle \otimes |\uparrow\rangle + |\chi\rangle \otimes |\downarrow\rangle\,, \tag{2.58}$$

given in the position representation by

$$\langle \mathbf{x}_{\rm e}|\Psi\rangle = \phi(\mathbf{x}_{\rm e})|\uparrow\rangle + \chi(\mathbf{x}_{\rm e})|\downarrow\rangle \tag{2.59}$$

involving a *pair* of wavefunctions ϕ, $\chi \in L^2(\mathbb{R}^3, d^3x)$. We'll investigate this in detail in section 5.3.

Let's now understand how operators act on our composite system. Given linear operators $A : \mathcal{H}_1 \to \mathcal{H}_1$ and $B : \mathcal{H}_2 \to \mathcal{H}_2$, we define the linear operator $A \otimes B : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$(A \otimes B)\,:\,|e_a\rangle\otimes|f_\alpha\rangle \;\mapsto\; (A|e_a\rangle) \otimes (B|f_\alpha\rangle) \tag{2.60}$$

and we extend this definition to arbitrary states in $\mathcal{H}_1 \otimes \mathcal{H}_2$ by linearity. In particular, the operator A acting purely on $\mathcal{H}_1$ corresponds to the operator $A \otimes 1_{\mathcal{H}_2}$ when acting on the tensor product, and similarly $1_{\mathcal{H}_1} \otimes B$ acts non–trivially just on the second factor in the tensor product. Note that since they act on separate Hilbert spaces,

$$[\,A \otimes 1_{\mathcal{H}_2}\,, 1_{\mathcal{H}_1} \otimes B\,] = 0 \tag{2.61}$$

for any operators A, B, even if these operators happen not to commute when acting on the same Hilbert space.

As an example, let's again consider the case of the Hydrogen atom, where the Hamiltonian takes the form

$$H = \frac{\mathbf{P}_{\rm e}^2}{2m_{\rm e}} \otimes 1_{\rm p} + 1_{\rm e} \otimes \frac{\mathbf{P}_{\rm p}^2}{2m_{\rm p}} + V(\mathbf{X}_{\rm e}, \mathbf{X}_{\rm p})\,. \tag{2.62}$$

The kinetic terms for the electron and proton each act non–trivially only on one factor in the tensor product, while the Coulomb interaction

$$V(\mathbf{X}_{\rm e}, \mathbf{X}_{\rm p}) = -\frac{e^2}{|\mathbf{X}_{\rm e} - \mathbf{X}_{\rm p}|} \tag{2.63}$$

depends non–trivially on the location of each particle. In this case, since V depends only on the relative separation of the two, it's actually more convenient to view the tensor product differently, writing

$$\mathcal{H}_{\rm Hyd} = \mathcal{H}_{\rm e} \otimes \mathcal{H}_{\rm p} = \mathcal{H}_{\rm com} \otimes \mathcal{H}_{\rm rel} \tag{2.64}$$

to split it in terms of the states describing the behaviour of the centre of mass and those describing the relative states. Defining the centre of mass and relative operators

$$\begin{aligned} \mathbf{X}_{\mathrm{com}} &= \frac{m_{\mathrm{e}}\mathbf{X}_{\mathrm{e}} + m_{\mathrm{p}}\mathbf{X}_{\mathrm{p}}}{m_{\mathrm{e}} + m_{\mathrm{p}}} & \qquad \mathbf{P}_{\mathrm{com}} &= \mathbf{P}_{\mathrm{e}} + \mathbf{P}_{\mathrm{p}} \\ \mathbf{X} &= \mathbf{X}_{\mathrm{e}} - \mathbf{X}_{\mathrm{p}} & \qquad \mathbf{P} &= \frac{m_{\mathrm{p}}\mathbf{P}_{\mathrm{e}} - m_{\mathrm{e}}\mathbf{P}_{\mathrm{p}}}{m_{\mathrm{e}} + m_{\mathrm{p}}}\,, \end{aligned} \tag{2.65}$$

the Hamiltonian becomes

$$H = \frac{\mathbf{P}_{\mathrm{com}}^2}{2M} \otimes 1_{\mathrm{rel}} + 1_{\mathrm{com}} \otimes \left[\frac{\mathbf{P}^2}{2\mu} - \frac{e^2}{|\mathbf{X}|}\right] \tag{2.66}$$

where $M = m_{\mathrm{e}} + m_{\mathrm{p}}$ is the total mass and $\mu = m_{\mathrm{e}}m_{\mathrm{p}}/M$ the *reduced mass*. Such calculations should be familiar from studying planetary orbits in IA Dynamics & Relativity.

Henceforth we'll often omit the ⊗ symbol — both in states and in operators — when the meaning is clear.

2.4 Postulates of Quantum Mechanics

So far, we've just been doing linear algebra, talking about Hilbert spaces in a fairly abstract way. Let's now begin to connect this to some physics.

The first postulate of quantum mechanics says that our system is described (up to a redundancy discussed below) by some state $|\psi\rangle \in \mathcal{H}$, and that any complete set of orthogonal states $\{|\phi_1\rangle, |\phi_2\rangle, \ldots\}$ is in one-to-one correspondence with all the possible outcomes of the measurement of some quantity. Further, if a system is prepared to be in some general state

$$|\psi\rangle = \sum_a c_a |\phi_a\rangle \tag{2.67}$$

then the *probability* that the measurement will yield an outcome corresponding to the state $|\phi_b\rangle$ is

$$\mathrm{Prob}(|\psi\rangle \to |\phi_b\rangle) = \frac{|\langle\phi_b|\psi\rangle|^2}{\langle\psi|\psi\rangle\,\langle\phi_b|\phi_b\rangle} \tag{2.68}$$

or equivalently $\mathrm{Prob}(|\psi\rangle \to |\phi_b\rangle) = |c_b|^2\,\langle\phi_b|\phi_b\rangle\,/\,\langle\psi|\psi\rangle$. The denominators in (2.68) allow for the possibility that the states $|\psi\rangle$ and $|\phi_b\rangle$ may have arbitrary norm. Notice that $\mathrm{Prob}(|\psi\rangle \to |\phi_b\rangle) \geq 0$ and also, since $\langle\psi|\psi\rangle = \sum_{a,b} c_a\overline{c_b}\langle\phi_b|\phi_a\rangle = \sum_b |c_b|^2\langle\phi_b|\phi_b\rangle$ by the orthogonality of the $|\phi_b\rangle$s, that

$$\sum_b \mathrm{Prob}(|\psi\rangle \to |\phi_b\rangle) = \sum_b \frac{|c_b|^2\langle\phi_b|\phi_b\rangle}{\langle\psi|\psi\rangle} = \frac{\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle} = 1\,. \tag{2.69}$$

Thus the probabilities sum to one.

Let me make some remarks. Firstly, note that we've not specified exactly which Hilbert space $\mathcal{H}$ we should use; is $\mathcal{H}$ to be $\mathbb{C}^n$, or ℓ^2, or $L^2(\mathbb{R}^3, d^3x)$, or something else? In fact, the appropriate choice depends on the system we're trying to describe. For example, a single point particle with no internal structure could be described using[16] $\mathcal{H} \cong L^2(\mathbb{R}^3, d^3x)$ whereas to specify the state of a more complicated system with several degrees of freedom we'll need a larger Hilbert space, encoding for example the location of its centre of mass, but also details of the system's orientation. We'll consider how to do this in chapter 2.3.

Second, the fact that quantum mechanics only predicts the probabilities of obtaning a given experimental outcome – even *in principle* – is one of its most puzzling features. Originally, this probabilistic interpretation of wavefunctions was due to Max Born, and (2.68) is sometimes known as the *Born rule.* He found that, solving the Schrödinger equation for scattering a particle off a generic obstacle in $\mathbb{R}^3$, the particle's wavefunction would typically become very spread out so as to have appreciable magnitude over a wide region. This is in contrast to our experience of particles bouncing off targets, each in a certain specific way. For example, recall that when performing his gold foil experiment to probe the structure of the atom, Rutherford usually observed α-particles to plough straight through, but occasionally saw them rebound back revealing the presence of the dense nucleus. Nucleus or not, each α-particle was observed at a specific angle from the target, and did not itself 'spread out'. Reconciling this observation with the behaviour of the wavefunction is what led Born to suggest that quantum mechanics should be inherently probabilistic. His reasoning was perhaps not totally sound, and we'll explore the interpretative meaning of QM further and from a more modern perspective in chapter 11.

As a final remark on this postulate, to clean up the basic probability rule (2.68), it's usually convenient to assume our states are *normalised*, meaning that

$$\langle\psi|\psi\rangle = 1\,, \tag{2.70}$$

and to expand them in an orthonormal basis $\{|\phi_b\rangle\}$. In this case, the above probability is simply

$$\mathrm{Prob}(|\psi\rangle \to |\phi_b\rangle) = |\langle\phi_b|\psi\rangle|^2\,, \tag{2.71}$$

while the quantity $\langle\phi_b|\psi\rangle \in \mathbb{C}$ itself is often called the *probability amplitude*, or just *amplitude* for short. In what follows, we'll almost always use this simpler expression, but it's important to recall that it only holds in the case of a properly normalised state expanded in an orthonormal basis. Even when all states are properly normalised, we're still free to redefine $|\psi\rangle \to \mathrm{e}^{\mathrm{i}\alpha}|\phi\rangle$ for some constant phase α. This phase drops out of both the normalisation condition and the Born rule (2.71) for the probability. Taking into account both the normalisation condition and the phase freedom, in quantum mechanics physical states do not correspond to elements $|\psi\rangle \in \mathcal{H}$ but rather to *rays through the origin.* That is, provided $|\psi\rangle \neq \mathbf{0}$, the entire family of states

$$|\psi_\lambda\rangle = \lambda|\psi\rangle \qquad \text{for } \lambda \in \mathbb{C}^* \tag{2.72}$$

define the same physical system — by tuning $|\lambda| = 1/\langle\psi|\psi\rangle$ we can always find a member this family that's correctly normalised, and the (constant) phase of λ cannot affect the

---

[16] Note that we only require $\mathcal{H}$ to be *isomorphic* to the space $L^2(\mathbb{R}^3, d^3x)$ of normalizable position-space wavefunctions. We could equally well describe our structureless particle using a momentum space wavefunction, and indeed $L^2(\widetilde{\mathbb{R}}^3, d^3p) \cong L^2(\mathbb{R}^3, d^3x)$, with the isomorphism provided by the Fourier transform.

probability of any outcome of a physical experiment. (This is the redundancy referred to above.) Note in particular that the zero vector $\mathbf{0}$ – the unique element of $\mathcal{H}$ with vanishing norm – never represents a physical state. Geometrically, the equivalence $|\psi\rangle \sim \lambda|\psi\rangle$ for all $|\psi\rangle \in \mathcal{H}/\{\mathbf{0}\}$ and all $\lambda \in \mathbb{C}^*$ means that physical states actually correspond to elements of the *projective* Hilbert space $\mathbb{P}\mathcal{H}$. As with any projective space, it's often most convenient to work simply with normalised vectors $|\psi\rangle \in \mathcal{H}$ themselves, and recall that the overall phase is irrelevant.

The second postulate of quantum mechanics states that observable quantities are represented by Hermitian linear operators. In particular, upon measurement of a quantity corresponding to a Hermitian operator Q, a state is certain to return the definite value q iff it is an eigenstate of Q with eigenvalue q. Let $|n\rangle$ be a complete, orthonormal set of eigenstates of some Hermitian operator Q, with $Q|n\rangle = q_n|n\rangle$. Then we can expand an arbitrary state $|\psi\rangle$ in this basis as $|\psi\rangle = \sum_n c_n|n\rangle$. The *expectation value* of Q in this state is

$$\langle Q\rangle_\psi = \langle\psi|Q|\psi\rangle = \sum_{m,n} \overline{c_m}\, c_n \langle m|Q|n\rangle = \sum_n q_n |c_n|^2\,, \tag{2.73}$$

using the orthonormality of the basis. This is just the sum of values of Q possessed by the states $|n\rangle$, weighted by the probability that $|\psi\rangle$ agrees with $|n\rangle$.

Since operators representing observables are Hermitian, for any such operator we have

$$\langle\psi|Q^2|\psi\rangle = \langle\psi|Q^\dagger Q|\psi\rangle = \|Q|\psi\rangle\|^2 \geq 0 \tag{2.74}$$

and hence

$$0 \leq \langle\psi|(Q - \langle Q\rangle_\psi)^2|\psi\rangle = \langle Q^2\rangle_\psi - \langle Q\rangle_\psi^2\,. \tag{2.75}$$

This shows that $\langle Q^2\rangle_\psi \geq \langle Q\rangle_\psi^2$, with equality iff $|\psi\rangle$ is an eigenstate of Q. We define the *rms deviation* $\Delta_\psi Q$ of Q from its mean $\langle Q\rangle_\psi$ in a state $|\psi\rangle$ by

$$\Delta_\psi Q = \sqrt{\langle\psi|(Q - \langle Q\rangle_\psi)^2|\psi\rangle}\,. \tag{2.76}$$

This is just the usual definition familiar from probability. As always, it gives us a measure of how 'spread' a state is around the eigenstate of Q. This implies that we can be sure of the value we'll obtain when measuring an observable quantity only if we somehow know that our state is in an eigenstate of the corresponding operator before carrying out the measurement.

Let me emphasize that these two postulates do not say anything about how the physical process of actually carrying out a measurement is described in the formalism of QM. (Nor do they even tell us what constitutes 'makes a measurement'.) In particular, we do *not* say that measuring the observable corresponding to some Hermitian operator Q has anything to do with the mathematical operation of acting on our state $|\psi\rangle$ with Q. According to the Copenhagen interpretation, if we measure the observable corresponding Q and find the result q, then immediately *after* this measurement, our system must be the state $|q\rangle$, because we've just learned that it does indeed have this value. The wavefunction is assumed to have *collapsed* from whatever it was before we measured it to $|q\rangle$. The Copenhagen interpretation is the most widely accepted version of quantum mechanics, but it's not uncontroversial. We'll try to understand measurement from a deeper perspective in section 11.6.

The final postulate of quantum mechanics is that the state $|\psi\rangle$ of our system evolves in time according to the *Schrödinger equation*

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle \tag{2.77}$$

where H is some distinguished operator known as the *Hamiltonian.*

2.4.1 The Generalized Uncertainty Principle

Suppose we have two Hermitian operators A and B. Let's define $|\psi_A\rangle$ to be the state $A|\psi\rangle - \langle A\rangle_\psi|\psi\rangle$ so that $(\Delta_\psi A)^2 = \||\psi_A\rangle\|^2$, and similarly $|\psi_B\rangle = B|\psi\rangle - \langle B\rangle_\psi|\psi\rangle$. Then the Cauchy–Schwarz inequality says

$$(\Delta_\psi A)^2(\Delta_\psi B)^2 = \|\,|\psi_A\rangle\|^2\,\|\,|\psi_B\rangle\|^2 \geq |\langle\psi_A|\psi_B\rangle|^2\,. \tag{2.78}$$

Expanding the *rhs*, we have

$$\langle\psi_A|\psi_B\rangle = \langle\psi|(A-\langle A\rangle_\psi)(B-\langle B\rangle_\psi)|\psi\rangle = \langle\psi|AB - \langle A\rangle_\psi\langle B\rangle_\psi|\psi\rangle \tag{2.79}$$

Since we're considering Hermitian opertaors A, B, we have $\langle\psi|AB|\psi\rangle = \overline{\langle\psi|BA|\psi\rangle}$, so

$$\mathrm{Im}\,\langle\psi_A|\psi_B\rangle = \frac{1}{2\mathrm{i}}\langle\psi|[A,B]|\psi\rangle\,. \tag{2.80}$$

Combining this with the Cauchy–Schwarz inequality gives the *generalised uncertainty relation*

$$(\Delta_\psi A)^2(\Delta_\psi B)^2 \geq |\langle\psi_A|\psi_B\rangle|^2 \geq \frac{1}{4}|\langle[A,B]\rangle_\psi|^2\,. \tag{2.81}$$

In particular, if $[A,B] \neq 0$ we cannot in general find states that simultaneously have definite values for both the quantities represented by A or B. As a particular case, recall from IB QM that the position and momentum operators obey the commutation relation $[X,P] = \mathrm{i}\hbar$. Using this in (2.81) gives

$$\Delta_\psi X\,\Delta_\psi P \geq \frac{\hbar}{2} \tag{2.82}$$

so that no quantum state can have a definite value of both position and momentum[17]. This is the original uncertainty principle of Heisenberg.

---

[17] The inequality is in fact saturated by states whose position space wavefunctions are Gaussians, so despite the fact that our derivation above neglected several positive semi–definite terms, we cannot place a higher bound on the minimum uncertainty.