

# **Principles of Quantum Mechanics**

**University of Cambridge Part II Mathematical Tripos**

---

**David Skinner**

*Department of Applied Mathematics and Theoretical Physics,  
Centre for Mathematical Sciences,  
Wilberforce Road,  
Cambridge CB3 0WA  
United Kingdom*

[d.b.skinner@damtp.cam.ac.uk](mailto:d.b.skinner@damtp.cam.ac.uk)

<http://www.damtp.cam.ac.uk/people/dbs26/>

**ABSTRACT:** These are the lecture notes for the Principles of Quantum Mechanics course given to students taking Part II Maths in Cambridge during Michaelmas Term of 2021. The main aim is to discuss quantum mechanics in the framework of Hilbert space, following Dirac. Along the way, we talk about transformations and symmetries, angular momentum, composite systems, dynamical symmetries, perturbation theory (both time-independent and time-dependent, degenerate and non-degenerate). We'll finish with a discussion of various interpretations of quantum mechanics, entanglement and decoherence.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Hilbert Space</b>	<b>4</b>
2.1	Definition of Hilbert Space	4
2.1.1	Examples	5
2.1.2	Dual Spaces	7
2.1.3	Dirac Notation and Continuum States	8
2.2	Operators	10
2.3	Composite Systems	14
2.3.1	Tensor Product of Hilbert Spaces	14
2.4	Postulates of Quantum Mechanics	18
2.4.1	The Generalized Uncertainty Principle	21
<b>3</b>	<b>The Harmonic Oscillator</b>	<b>22</b>
3.1	Raising and Lowering Operators	22
3.2	Dynamics of Oscillators	25
3.2.1	Coherent States	27
3.2.2	Anharmonic Oscillations	27
<b>4</b>	<b>Transformations and Symmetries</b>	<b>30</b>
4.1	Transformations of States and Operators	30
4.1.1	Continuous Transformations	32
4.2	Translations	33
4.3	Rotations	37
4.3.1	Translations Around a Circle	41
4.3.2	Spin	42
4.4	Time Translations	44
4.4.1	The Heisenberg Picture	45
4.5	Dynamics	46
4.5.1	Symmetries and Conservation Laws	49
4.6	Parity	50
<b>5</b>	<b>Angular Momentum</b>	<b>55</b>
5.1	Angular Momentum Eigenstates ( <i>A Little Representation Theory</i> )	56
5.2	Rotations and Orientation	58
5.2.1	Rotation of Diatomic Molecules	59
5.3	Spin	61
5.3.1	Large Rotations	61
5.3.2	The Stern–Gerlach Experiment	62
5.3.3	Spinors and Projective Representations	64

5.3.4	Spin Matrices	67
5.3.5	Paramagnetic Resonance and MRI Scanners	70
5.4	Orbital Angular Momentum	73
5.4.1	Spherical Harmonics	73
<b>6</b>	<b>Dynamical Symmetries</b>	<b>77</b>
6.1	The Three-Dimensional Isotropic Harmonic Oscillator	77
6.1.1	Treatment in Cartesian coordinates	78
6.2	Dynamical Symmetry of the $3d$ Isotropic Harmonic Oscillator	79
6.2.1	Angular Momentum of the Isotropic Oscillator	81
6.3	The Runge–Lenz Vector in Hydrogen	83
6.4	The Radial Hamiltonian for the Coulomb Potential	85
<b>7</b>	<b>Addition of Angular Momentum</b>	<b>89</b>
7.1	Combining the Angular Momenta of Two States	90
7.1.1	$j \otimes 0 = j$	94
7.1.2	$\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$	94
7.1.3	$1 \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2}$	95
7.1.4	The Classical Limit	96
7.2	Angular Momentum of Operators	97
7.2.1	The Wigner–Eckart Theorem	97
7.2.2	Dipole Moment Transitions	97
<b>8</b>	<b>Identical Particles</b>	<b>99</b>
8.1	Bosons and Fermions	99
8.1.1	Pauli’s Exclusion Principle	101
8.1.2	The Periodic Table	102
8.1.3	White Dwarfs, Neutron Stars and Supernovae	106
8.2	Exchange and Parity in the Centre of Momentum Frame	106
8.2.1	Identical Particles and Inelastic Collisions	106
<b>9</b>	<b>Perturbation Theory I: Time Independent Case</b>	<b>109</b>
9.1	An Analytic Expansion	109
9.1.1	Fine Structure of Hydrogen	112
9.1.2	Hyperfine Structure of Hydrogen	116
9.1.3	The Ground State of Helium	117
9.1.4	The Quadratic Stark Effect	120
9.2	Degenerate Perturbation Theory	122
9.2.1	The Linear Stark Effect	125
9.3	Does Perturbation Theory Converge?	126

<b>10 Perturbation Theory II: Time Dependent Case</b>	<b>129</b>
10.1 The Interaction Picture	129
10.2 Prodding the Harmonic Oscillator	132
10.3 A Constant Perturbation Turned On at $t = 0$	133
10.4 Fermi's Golden Rule	134
10.4.1 The Photoelectric Effect	136
10.4.2 Absorption and Stimulated Emission	139
10.4.3 Spontaneous Emission	141
<b>11 Interpreting Quantum Mechanics</b>	<b>145</b>
11.1 The Density Operator	145
11.1.1 The Bloch Sphere	147
11.2 Entropy	148
11.2.1 The Gibbs Distribution	150
11.3 Reduced Density Operators	150
11.4 Decoherence	151
11.5 Time Evolution of Density Operators and Reduced Density Operators	152
11.5.1 Decoherence and Measurement	154
11.6 Quantum Mechanics or Hidden Variables?	156
11.6.1 Bell's Inequality	158
11.6.2 The CHSH Inequality	160

---

## Acknowledgments

Nothing in these lecture notes is original. In particular, my treatment is heavily influenced by the textbooks listed below, especially Weinberg's *Lectures on Quantum Mechanics*, Hall's *Quantum Theory for Mathematicians* and Binney & Skinner's *The Physics of Quantum Mechanics* (though I feel more justified in plagiarising this one). I've also borrowed from previous versions of the lecture notes for this course, by Profs. R. Horgan and A. Davis, which are available online.

I am supported by the European Union under an FP7 Marie Curie Career Integration Grant.

## Preliminaries

These are the notes for the Part II course on Principles of Quantum Mechanics offered in Part II of the Maths Tripos, so I'll feel free to assume you took the IB courses on *Quantum Mechanics*, *Methods* and *Linear Algebra* last year. For Part II, I'd expect the material presented here to complement the courses on *Classical Dynamics*, *Statistical Physics* and perhaps *Integrable Systems* very well on the Applied side, in addition to being a prerequisite for *Applications of Quantum Mechanics* next term. On the pure side, the courses *Linear Analysis* or *Functional Analysis* and *Representation Theory* are probably the most closely related to this one.

## Books & Other Resources

There are many textbooks and reference books available on Quantum Mechanics. Different ones emphasise different aspects of the theory, or different applications in physics, or give prominence to different mathematical structures. QM is such a huge subject nowadays that it is probably impossible for a single textbook to give an encyclopaedic treatment (and *absolutely* impossible for me to do so in a course of 24 lectures). Here are some of the ones I've found useful while preparing these notes; you might prefer different ones to me.

- **Weinberg, S.** *Lectures on Quantum Mechanics*, CUP (2013).  
This is perhaps the single most appropriate book for the course. Written by one of the great physicists of our times, this book contains a wealth of information. Highly recommended.
- **Binney, J.J. and Skinner, D.** *The Physics of Quantum Mechanics*, OUP (2014).  
The number 1 international bestseller... Contains a much fuller version of these notes (with better diagrams!). Put it on your Christmas list.
- **Dirac, P.** *Principles of Quantum Mechanics*, OUP (1967).  
The notes from an earlier version of this course! Written with exceptional clarity and insight, this is a genuine classic of theoretical physics by one of the founders of Quantum Mechanics.
- **Messiah, A.** *Quantum Mechanics*, Vols 1 & 2, Dover (2014).  
Another classic textbook, originally from 1958. It provides a comprehensive treatment of QM, though the order of the presentation is slightly different to the one we'll follow in this course.
- **Shankar, R.** *Principles of Quantum Mechanics*, Springer (1994).  
A much-loved textbook, with particular emphasis on the physical applications of quantum mechanics.
- **Hall, B.** *Quantum Theory for Mathematicians*, Springer (2013).  
If you're more interested in the functional analysis & representation theory aspects of QM rather than the physical applications, this could be the book for you. Much of it is at a more advanced level than we'll need.

- **Sakurai, J.J. and Napolitano, J.** *Modern Quantum Mechanics*, CUP (3<sup>rd</sup> edition, 2020).

By now, another classic book on QM, presented from a perspective very close to this course. Also contains an excellent introduction to path integrals.

- **Townsend, J.** *A Modern Approach to Quantum Mechanics*, 2<sup>nd</sup> edition, University Science Books (2012).

Very accessible, but also fairly comprehensive textbook on QM. Starts by considering spin- $\frac{1}{2}$  particle as a simple example of a two-state system.

Textbooks are expensive, so check what's available in your college library before shelling out – any reasonably modern textbook on QM will be adequate for this course. There are also lots of excellent resources available freely online, including these:

- **Here** are the lecture notes from another recent version of this course.
- **Here** are the lecture notes from the Part II Further Quantum Mechanics course in the Cavendish.
- **These** are the lecture notes from Prof. W. Taylor's graduate course on QM at MIT. The first half of his course covers material that is relevant here, while scattering theory will be covered in next term's Applications of Quantum Mechanics Part II course.
- **This** is the lecture course on Quantum Mechanics from Prof. R. Fitzpatrick at the University of Texas, Austin.

## 1 Introduction

In classical mechanics, a particle's motion is governed by Newton's Laws. These are second order o.d.e.s, so to determine the fate of our particle we must specify two initial conditions. We could take these to be the particle's initial position  $\mathbf{x}(t_0)$  and velocity  $\mathbf{v}(t_0)$ , or it's initial position  $\mathbf{x}(t_0)$  and momentum  $\mathbf{p}(t_0)$ . Once these are specified the motion is determined (provided of course we understand how to describe the forces that are acting). This means that we can solve Newton's Second Law to find the values  $(\mathbf{x}(t), \mathbf{p}(t))$  for  $t > t_0$ <sup>1</sup>. So as time passes, our classical particle traces out a trajectory in the space  $M$  of possible positions and momenta, sketched in figure 1. The space  $M$  is known as *phase space* and in our case, for motion in three dimensions,  $M$  is just  $\mathbb{R}^6$ . In general  $M$  comes with a rich geometry known as a *Poisson structure*; you'll study this structure in detail if you're taking the Part II courses on Classical Dynamics or Integrable Systems, and we'll touch on it later in this course, too.

Classical observables are represented by functions

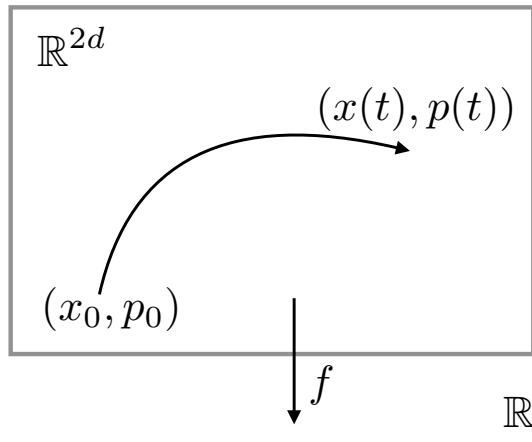
$$f : M \rightarrow \mathbb{R} \quad f : (\mathbf{x}, \mathbf{p}) \mapsto f(\mathbf{x}, \mathbf{p}).$$

For example, we may be interested in the kinetic energy  $T = \mathbf{p}^2/2m$ , potential energy  $V(\mathbf{x})$ , angular momentum  $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ , or a host of other possible quantities. A priori, these functions are defined everywhere over  $M$ , but if we want to know about the energy or angular momentum of *our specific particle* then we should evaluate them not at some random point  $(\mathbf{x}, \mathbf{p}) \in M$ , but *along the particle's phase space trajectory*. For example, if at time  $t$  the particle has position  $\mathbf{x}(t)$  and momentum  $\mathbf{p}(t)$ , then its angular momentum is  $\mathbf{x}(t) \times \mathbf{p}(t)$ . Thus the values of the particle's energy, angular momentum *etc.* may depend on time, though of course our *definition* of these quantities does not<sup>2</sup>. In this way,

---

<sup>1</sup>We can certainly do this numerically, at least for  $t$  in a sufficiently small open subset around  $t_0$ .

<sup>2</sup>Technically, we take the pullback of  $f$  by the embedding map  $\iota : [t_0, \infty) \rightarrow M$ .



**Figure 1:** A particle's trajectory in phase space. Observables are represented by functions  $f : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ , evaluated along a given particle's trajectory.

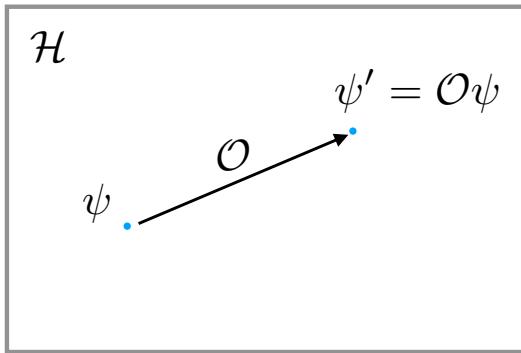
everything we could possibly want to know about a single, pointlike particle is encoded in its phase space trajectory.

But that is not our World. To the best of our current experimental knowledge, our World is a quantum, not classical, one. Initially, these experiments were based on careful studies of atomic spectroscopy and blackbody radiation, but nowadays I'd prefer to say that the best evidence of quantum mechanics is simply that we use it constantly in our everyday lives. Each time you listen to music on your stereo, post a photo on Instagram or make a call on your phone you're relying on technology that's only become possible due to our understanding of the quantum structure of matter. Whenever you plug something into the mains, you're using electricity that's in part generated by nuclear reactions in which quantum mechanics is essential, while much of modern medicine relies on new drugs designed with the benefit of the improved understanding of chemistry that quantum mechanics provides.

In such a quantum world, instead of a phase space trajectory, everything we could want to know about a particle is encoded in a vector  $\psi$  in Hilbert space  $\mathcal{H}$ . As you met in IB Quantum Mechanics, this *state vector* evolves in time according to Schrödinger's equation. In the quantum world, observables are represented by certain operators  $\mathcal{O}$ . The operators you saw in IB QM had the same sort of form as observables classical mechanics, such as the kinetic energy operator  $\hat{T} = \hat{\mathbf{p}}^2/2m$  or angular momentum operator  $\hat{\mathbf{L}} = \hat{\mathbf{x}} \times \hat{\mathbf{p}}$ . However, rather than being functions, these operators are (roughly) linear maps

$$\mathcal{O} : \mathcal{H} \rightarrow \mathcal{H}.$$

Again, in the first instance these operators are defined throughout  $\mathcal{H}$ , but if we're interested in knowing about the energy or angular momentum of our particular quantum particle, then we should find out what happens when they act on the specific  $\psi \in \mathcal{H}$  that describes the state of our particle at time  $t$ . (See figure 2.)



**Figure 2:** In Quantum Mechanics, complete knowledge of a particle's states is determined by a vector in Hilbert space. Observables are represented by Hermitian linear operators  $\mathcal{O} : \mathcal{H} \rightarrow \mathcal{H}$ .

In the following chapters we'll study what Hilbert space is and what it's operators do in a more general framework than you saw last year, building your insight into the mathematical structure of quantum mechanics. Much of this is just linear algebra, but the Hilbert spaces we'll care about in QM are often infinite-dimensional, so we also make contact with Functional Analysis. Furthermore, although last year you 'guessed' the form of quantum operators by analogy with their classical counterparts, we'll see that at a deeper level many of them can be understood to have their origins in symmetries of space and time; the operators just reflect the way these symmetry transformations act on Hilbert space, rather than on (non-relativistic) space-time. In this way, Quantum Mechanics makes contact with Representation Theory.

So, mathematically, much of Quantum Mechanics boils down to a mix of Functional Analysis and Representation Theory. It's even true that it provides a particularly interesting example of these subjects. *But this is not the reason we study it.* We study Quantum Mechanics in an effort to understand and appreciate *our* World, not some abstract mathematical one. You're all intimately familiar with vector spaces, and you're (hopefully!) also very good at solving Sturm–Liouville type eigenfunction / eigenvalue problems. But the real skill is in understanding how this formalism relates to the world we see around us.

It's not obvious. Newton's laws are (at least generically) non-linear differential equations and we can't usually superpose solutions. General Relativity teaches us that space-time is not flat. So it's not at all clear that our particle should in fact be described by a point in a vector space, any more than it was obvious to Aristotle that bodies actually stay in uniform motion unless acted on by a force, or clear to the Ancients that the arrival of solar eclipses, changes of the weather, or any other natural phenomenon are actually governed by calculable Laws, rather than the whims of various gods. For this reason, instead of emphasising how weird and different Quantum Mechanics is, I'd prefer to make you appreciate how it actually underpins the physics you're already familiar with. No matter how good you are at solving eigenvalue problems, if you don't see how these relate to your everyday physical intuition, knowledge you've built up since first opening your eyes and learning to crawl, then you haven't really understood the subject.

Let's begin.

## 2 Hilbert Space

The realm of Quantum Mechanics is Hilbert space<sup>3</sup>, 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

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

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  $a, b, c, \dots \in \mathbb{C}$ , called scalars. This multiplication is

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

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

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

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

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

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

---

<sup>3</sup>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.

<sup>4</sup>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.

<sup>5</sup>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  $\{\phi_1, \phi_2, \dots, \phi_n\}$  are *linearly independent* iff the only solution to

$$c_1\phi_1 + c_2\phi_2 + \dots + c_n\phi_n = o \quad (2.5)$$

for  $c_i \in \mathbb{C}$  is  $c_1 = c_2 = \dots = c_n = 0$ . 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  $\{\phi_1, \phi_2, \dots, \phi_n\}$  is *orthonormal* with respect to  $(\cdot, \cdot)$  if

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

An orthonormal set  $\{\phi_1, \phi_2, \dots, \phi_n\}$  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  $\phi_b$ , since

$$(\phi_b, \psi) = \left( \phi_b, \sum_a c_a \phi_a \right) = \sum_a c_a (\phi_b, \phi_a) = c_b \quad (2.7)$$

using the linearity of the inner product and orthonormality of the  $\phi_a$ .

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  $\{S_1, S_2, \dots\}$  *converges in the norm* if there is some vector  $\Psi \in \mathcal{H}$  to which it converges, in the sense that

$$\lim_{N \rightarrow \infty} \|S_N - \Psi\| = 0. \quad (2.8)$$

To say that  $\mathcal{H}$  is *complete* (or complete in the norm (2.4)) means that every Cauchy sequence  $\{S_1, S_2, \dots\}$  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 analyse: 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

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

and corresponding norm

$$\|\mathbf{u}\| = \sqrt{\sum_{i=1}^n |u_i|^2}. \quad (2.10)$$

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  $\mathbf{u} = (u_1, u_2, u_3, \dots)$  such that

$$\sum_{i=1}^{\infty} |u_i|^2 < \infty. \quad (2.11)$$

This space is known as  $\ell^2$  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

$$(\mathbf{v}, \mathbf{u}) = \sum_{i=1}^{\infty} \bar{v}_i u_i \quad (2.12)$$

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  $\ell^2$  with respect to this norm enables us to meaningfully take limits and, ultimately, differentiate vectors  $\mathbf{u} \in \ell^2$ .

In physics,  $\ell^2$  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  $\ell^2$ . Similarly, the space of *bound* states ( $E < 0$ ) of the hydrogen atom is  $\ell^2$ , 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  $\psi : \mathbb{R} \rightarrow \mathbb{C}$  and  $\phi : \mathbb{R} \rightarrow \mathbb{C}$ , we can define linear combinations  $a\psi + b\phi$  for any  $a, b \in \mathbb{C}$  in the obvious way

$$(a\psi + b\phi)(x) = a\psi(x) + b\phi(x) \quad (2.13)$$

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

$$\|\psi\| = \sqrt{\int_{\mathbb{R}} |\psi(x)|^2 dx} \quad (2.14)$$

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  $|x| \rightarrow \infty$ , 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<sup>6</sup>.

Finally, we take the inner product between two functions  $\psi, \phi$  whose norms are finite to be

$$(\phi, \psi) = \int_{\mathbb{R}} \bar{\phi}(x) \psi(x) dx \quad (2.15)$$

and again this converges by Cauchy–Schwarz. In physics, we often call  $\psi(x)$  the *wavefunction* of the particle, and sometimes call  $(\phi, \psi)$  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  $\mathcal{H} \rightarrow \mathbb{C}$ . 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

$$\varphi : a\psi_1 + b\psi_2 \mapsto a\varphi(\psi_1) + b\varphi(\psi_2) \quad (2.16)$$

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  $(\phi, \cdot)$  in  $\mathcal{H}^*$  which acts by

$$(\phi, \cdot) : \psi \mapsto (\phi, \psi) \quad (2.17)$$

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

---

<sup>6</sup>The requirement that the space be *complete* in the norm (2.14) is rather subtle. If  $\|\psi - \phi\| = 0$ , then we must identify  $\psi$  and  $\phi$  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  $L^2(\mathbb{R}, dx)$  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.)  $L^2(\mathbb{R}, dx)$  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  $\varphi : \mathcal{H} \rightarrow \mathbb{C}$  can be written as  $(\phi, \cdot)$  for some fixed choice of  $\phi \in \mathcal{H}$ . This means in particular that the inner product  $(\cdot, \cdot)$  provides a vector space isomorphism

$$\mathcal{H}^* \xrightarrow{(\cdot, \cdot)} \mathcal{H}. \quad (2.18)$$

Comfortingly, the same result also holds in infinite dimensions<sup>7</sup>, 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 “|” is known as a *ket*. An element of the dual space is written  $\langle\phi|$  and the symbol “⟨” 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  $\phi$  vs  $(\phi, \cdot)$ . 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 *antilinear* 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

$$|\psi\rangle = \sum_a \psi_a |e_a\rangle \quad (2.19)$$

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

$$\langle a'|a\rangle = \delta(a' - a) \quad (2.20)$$

using the Dirac  $\delta$ -function. Then we write

$$|\psi\rangle = \int \psi(a) |a\rangle da \quad (2.21)$$

---

<sup>7</sup>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  $\varphi : \mathcal{H} \rightarrow \mathbb{C}$  from the *continuous dual*, where the number  $\varphi(\psi)$  is required to be vary continuously as  $\psi$  varies in  $\mathcal{H}$ . The Riesz Representation applies to the continuous dual. For example, the algebraic dual also contains distributions such as the Dirac  $\delta$ , 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  $\psi$ . Functional analysis is almost always interested in the continuous dual, and this is 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

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

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

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

we see that the complex coefficients are

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

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  $|\psi\rangle = \int_{\mathbb{R}} \tilde{\psi}(p) |p\rangle dp$ , where the new coefficients  $\tilde{\psi}(p) = \langle p | \psi \rangle$  are the *momentum space wavefunction*. Later, we’ll show that  $\langle x | p \rangle = e^{ixp/\hbar} / \sqrt{2\pi\hbar}$ , so these two sets of coefficients are related by

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

(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 position 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<sup>8</sup> 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

$$\|x\rangle\|^2 = \delta(x - x) = \delta(0) \quad ??? \quad (2.26)$$

---

<sup>8</sup>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<sup>9</sup>  $A : \mathcal{H} \rightarrow \mathcal{H}$  that is compatible with the vector space structure in the sense that

$$A(c_1|\phi_1\rangle + c_2|\phi_2\rangle) = c_1A|\phi_1\rangle + c_2A|\phi_2\rangle, \quad (2.27)$$

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

$$(\alpha A + \beta B) : |\phi\rangle \mapsto \alpha A|\phi\rangle + \beta B|\phi\rangle \quad (2.28)$$

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

$$AB : \phi \mapsto A \circ B|\phi\rangle = A(B|\phi\rangle) \quad (2.29)$$

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)<sup>10</sup>.

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*

$$[A, B] = AB - BA. \quad (2.30)$$

<sup>9</sup>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  $\psi(x) \in L^2(\mathbb{R}, dx)$  by  $-i\hbar \partial/\partial x$ . If  $\psi(x)$  is discontinuous, then it could be that  $\int |\psi'|^2 dx$  diverges even though  $\int |\psi|^2 dx$  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.

<sup>10</sup>A linear operator is *bounded* if, for all  $|\psi\rangle \in \mathcal{H}$ ,

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

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:

antisymmetry	$[A, B] = -[B, A]$	(2.31)
linearity	$[\alpha_1 A_1 + \alpha_2 A_2, B] = \alpha_1 [A_1, B] + \alpha_2 [A_2, B] \quad \forall \alpha_1, \alpha_2 \in \mathbb{C}$	
Leibniz identity	$[A, BC] = [A, B]C + B[A, C]$	
Jacobi identity	$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$	

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

$$A|\psi\rangle = a_\psi|\psi\rangle \quad (2.32)$$

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  $\delta$ -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

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

One can easily<sup>11</sup> check that

$$(A+B)^\dagger = A^\dagger + B^\dagger, \quad (AB)^\dagger = B^\dagger A^\dagger, \quad (\alpha A)^\dagger = \bar{\alpha} A^\dagger \quad \text{and} \quad (A^\dagger)^\dagger = A, \quad (2.34)$$

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

$$\langle a | A^\dagger = \langle a | \bar{a}, \quad (2.35)$$

---

<sup>11</sup>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*<sup>12</sup> 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

$$q\langle q|q\rangle = \langle q|Q|q\rangle = \overline{\langle q|Q|q\rangle} = \bar{q}\langle q|q\rangle, \quad (2.36)$$

so  $\bar{q} = q$  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

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

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

$$Q = \sum_n q_n |n\rangle\langle n| \quad (2.38)$$

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

$$Q|\psi\rangle = \sum_n q_n |n\rangle\langle n|\psi\rangle. \quad (2.39)$$

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

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

---

<sup>12</sup>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  $-i\hbar d/dx$  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  $\text{Dom}(Q) \subseteq \mathcal{H}$  of the operator is the same as the domain  $\text{Dom}(Q^\dagger)$  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

$$f(Q) = \sum_n f(q_n) |n\rangle\langle n| \quad (2.41)$$

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

$$Q^{-1} = \sum_n \frac{1}{q_n} |n\rangle\langle n| \quad (2.42)$$

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

$$\ln Q = \sum_n \ln(q_n) |n\rangle\langle n| \quad (2.43)$$

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

$$1_{\mathcal{H}} = \sum_n |n\rangle\langle n| \quad \text{or} \quad 1_{\mathcal{H}} = \int |q\rangle\langle q| dq \quad (2.44)$$

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

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

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

$$\langle x|P|\psi\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|\psi\rangle, \quad (2.46)$$

or  $\hat{P}\psi(x) = -i\hbar\partial\psi/\partial x$  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

$$A_{nm} := \langle n|A|m\rangle \quad (2.47)$$

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<sup>13</sup>.

### 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<sup>14</sup>. It follows that

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

---

<sup>13</sup>Though see the previous footnote for subtleties in the infinite dimensional case.

<sup>14</sup>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

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

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

$$|\psi_1\rangle = \sum_{a=1}^m c_a |e_a\rangle \quad \text{and} \quad |\psi_2\rangle = \sum_{\alpha=1}^n d_\alpha |f_\alpha\rangle, \quad (2.50)$$

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

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

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  $\langle \cdot | \cdot \rangle_1$  and  $\langle \cdot | \cdot \rangle_2$  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

$$(\langle e_a | \otimes \langle f_\alpha |) (|e_b\rangle \otimes |f_\beta\rangle) = \langle e_a | e_b \rangle_1 \langle f_\alpha | f_\beta \rangle_2, \quad (2.52)$$

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

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

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

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

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

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

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

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

$$|\Psi\rangle = \int_{\mathbb{R}^6} \Psi(\mathbf{x}_e, \mathbf{x}_p) |\mathbf{x}_e\rangle \otimes |\mathbf{x}_p\rangle d^3x_e d^3x_p \quad (2.57)$$

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

---

<sup>15</sup>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  $f : \mathbb{R}^2 \rightarrow \mathbb{C}$  given by

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

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  $\hat{\mathcal{H}}_1 \hat{\otimes} \mathcal{H}_2$ , 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  $\mathcal{H}_e = L^2(\mathbb{R}^3, d^3x) \otimes \mathbb{C}^2$ , 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

$$|\Psi\rangle = |\phi\rangle \otimes |\uparrow\rangle + |\chi\rangle \otimes |\downarrow\rangle, \quad (2.58)$$

given in the position representation by

$$\langle \mathbf{x}_e | \Psi \rangle = \phi(\mathbf{x}_e) | \uparrow \rangle + \chi(\mathbf{x}_e) | \downarrow \rangle \quad (2.59)$$

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

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

$$(A \otimes B) : |e_a\rangle \otimes |f_\alpha\rangle \mapsto (A|e_a\rangle) \otimes (B|f_\alpha\rangle) \quad (2.60)$$

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,

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

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

$$H = \frac{\mathbf{P}_e^2}{2m_e} \otimes 1_p + 1_e \otimes \frac{\mathbf{P}_p^2}{2m_p} + V(\mathbf{X}_e, \mathbf{X}_p). \quad (2.62)$$

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

$$V(\mathbf{X}_e, \mathbf{X}_p) = -\frac{e^2}{|\mathbf{X}_e - \mathbf{X}_p|} \quad (2.63)$$

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

$$\mathcal{H}_{\text{Hyd}} = \mathcal{H}_e \otimes \mathcal{H}_p = \mathcal{H}_{\text{com}} \otimes \mathcal{H}_{\text{rel}} \quad (2.64)$$

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

$$\begin{aligned}\mathbf{X}_{\text{com}} &= \frac{m_e \mathbf{X}_e + m_p \mathbf{X}_p}{m_e + m_p} & \mathbf{P}_{\text{com}} &= \mathbf{P}_e + \mathbf{P}_p \\ \mathbf{X} &= \mathbf{X}_e - \mathbf{X}_p & \mathbf{P} &= \frac{m_p \mathbf{P}_e - m_e \mathbf{P}_p}{m_e + m_p},\end{aligned}\tag{2.65}$$

the Hamiltonian becomes

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

where  $M = m_e + m_p$  is the total mass and  $\mu = m_e m_p / M$  the *reduced mass*. Such calculations should be familiar from studying planetary orbits in IA Dynamics & Relativity.

Henceforth we'll often omit the  $\otimes$  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  $\{|\phi_1\rangle, |\phi_2\rangle, \dots\}$  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

$$\text{Prob}(|\psi\rangle \rightarrow |\phi_b\rangle) = \frac{|\langle\phi_b|\psi\rangle|^2}{\langle\psi|\psi\rangle\langle\phi_b|\phi_b\rangle}\tag{2.68}$$

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

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

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  $\ell^2$ , 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<sup>16</sup>  $\mathcal{H} \cong L^2(\mathbb{R}^3, d^3x)$  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 obtaining 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  $\alpha$ -particles to plough straight through, but occasionally saw them rebound back revealing the presence of the dense nucleus. Nucleus or not, each  $\alpha$ -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

$$\langle \psi | \psi \rangle = 1, \quad (2.70)$$

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

$$\text{Prob}(|\psi\rangle \rightarrow |\phi_b\rangle) = |\langle \phi_b | \psi \rangle|^2, \quad (2.71)$$

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  $|\psi\rangle \rightarrow e^{i\alpha}|\phi\rangle$  for some constant phase  $\alpha$ . 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

$$|\psi_\lambda\rangle = \lambda|\psi\rangle \quad \text{for } \lambda \in \mathbb{C}^* \quad (2.72)$$

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  $\lambda$  cannot affect the

---

<sup>16</sup>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

$$\langle Q \rangle_\psi = \langle \psi | Q | \psi \rangle = \sum_{m,n} \bar{c}_m c_n \langle m | Q | n \rangle = \sum_n q_n |c_n|^2, \quad (2.73)$$

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

$$\langle \psi | Q^2 | \psi \rangle = \langle \psi | Q^\dagger Q | \psi \rangle = \|Q|\psi\rangle\|^2 \geq 0 \quad (2.74)$$

and hence

$$0 \leq \langle \psi | (Q - \langle Q \rangle_\psi)^2 | \psi \rangle = \langle Q^2 \rangle_\psi - \langle Q \rangle_\psi^2. \quad (2.75)$$

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

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

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*

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = H|\psi\rangle \quad (2.77)$$

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  $(\Delta_\psi A)^2 = \|\psi_A\|^2$ , and similarly  $|\psi_B\rangle = B|\psi\rangle - \langle B \rangle_\psi |\psi\rangle$ . Then the Cauchy–Schwarz inequality says

$$(\Delta_\psi A)^2 (\Delta_\psi B)^2 = \|\psi_A\|^2 \|\psi_B\|^2 \geq |\langle \psi_A | \psi_B \rangle|^2. \quad (2.78)$$

Expanding the *rhs*, we have

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

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

$$\text{Im } \langle \psi_A | \psi_B \rangle = \frac{1}{2i} \langle \psi | [A, B] | \psi \rangle. \quad (2.80)$$

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

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

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  $[X, P] = i\hbar$ . Using this in (2.81) gives

$$\Delta_\psi X \Delta_\psi P \geq \frac{\hbar}{2} \quad (2.82)$$

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

---

<sup>17</sup>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.

### 3 The Harmonic Oscillator

I now want to use Dirac's formalism to study a simple system – the one-dimensional harmonic oscillator – with which you should already be familiar. Our aim here is not to learn new things about harmonic oscillators; indeed, we'll mostly just recover results you've known about since you first heard of simple harmonic motion. Rather, our aim is to get accustomed to this more abstract approach to QM, seeing how it's a very powerful way to think about the subject.

The steps we follow in our treatment of the harmonic oscillator will form a good prototype for the way we'll approach many other problems in QM. We'll see these same steps repeated in various contexts throughout the course.

#### 3.1 Raising and Lowering Operators

The Hamiltonian of a harmonic oscillator of mass  $m$  and classical frequency  $\omega$  is

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2. \quad (3.1)$$

where  $X$  and  $P$  are the position and momentum operators, respectively. To analyse this, we begin by defining the dimensionless combination

$$A = \frac{1}{\sqrt{2m\hbar\omega}} (m\omega X + iP). \quad (3.2)$$

$A$  is not a Hermitian operator, but since  $X$  and  $P$  are both Hermitian, we see that the adjoint of  $A$  is

$$A^\dagger = \frac{1}{\sqrt{2m\hbar\omega}} (m\omega X - iP). \quad (3.3)$$

Roughly, the motivation for introducing these operators is that they allow us to ‘factorize’ the Hamiltonian. More precisely, we have

$$\begin{aligned} A^\dagger A &= \frac{1}{2m\hbar\omega} (m\omega X - iP)(m\omega X + iP) \\ &= \frac{1}{2m\hbar\omega} (P^2 + m^2\omega^2 X^2 + im\omega[X, P]) \\ &= \frac{H}{\hbar\omega} - \frac{1}{2} \end{aligned} \quad (3.4)$$

so we can write our Hamiltonian as

$$H = \hbar\omega \left( A^\dagger A + \frac{1}{2} \right) = \hbar\omega \left( N + \frac{1}{2} \right), \quad (3.5)$$

where

$$N = A^\dagger A. \quad (3.6)$$

$A$  and  $A^\dagger$  are often called *lowering* and *raising* operators, respectively, whilst  $N$  is often called the *number* operator. (The reason for these names will become apparent soon.) Notice that computing the spectrum of  $H$  is obviously equivalent to computing the spectrum of  $N$ .

Whenever we're presented with some new operators, the first thing to do is to work out their commutation relations. In this case, the fundamental commutation relations  $[X, P] = i\hbar$  show that

$$\begin{aligned} [A, A^\dagger] &= \frac{1}{2m\hbar\omega} (m^2\omega^2[X, X] - im\omega[X, P] + im\omega[P, X] + [P, P]) \\ &= -\frac{im\omega}{2m\hbar\omega} ([X, P] - [P, X]) \\ &= 1, \end{aligned} \tag{3.7}$$

whilst  $[A, A] = 0 = [A^\dagger, A^\dagger]$  trivially. It will also be useful to compute commutators involving  $N$ . We have

$$[N, A^\dagger] = [A^\dagger A, A^\dagger] = (A^\dagger [A, A^\dagger] + [A^\dagger, A^\dagger] A) = A^\dagger, \tag{3.8}$$

where the final step uses (3.7). We learn that, similarly,  $[N, A] = -A$  by taking the adjoint of (3.8).

Let's see what these commutators teach us about the possible energy levels. Suppose that  $|n\rangle$  is a correctly normalised eigenstate of  $N$ , so that  $N|n\rangle = n|n\rangle$  and  $\langle n|n\rangle = 1$ . We have

$$NA^\dagger|n\rangle = (A^\dagger N + [N, A^\dagger])|n\rangle = A^\dagger(N+1)|n\rangle = (n+1)A^\dagger|n\rangle \tag{3.9}$$

so  $A^\dagger|n\rangle$  is an eigenstate of  $N$  with eigenvalue  $n+1$ . Similarly,

$$NA|n\rangle = (AN + [N, A])|n\rangle = (n-1)A|n\rangle \tag{3.10}$$

so  $A|n\rangle$  is an eigenstate of  $N$  with eigenvalue  $n-1$ . Acting with  $A^\dagger$  thus raises the eigenvalue of  $N$  by one unit, whilst acting with  $A$  lowers it by one, giving these operators their names. The above algebra shows that if we can find just one energy eigenstate  $|n\rangle$ , then we can construct a whole family of them by repeatedly applying either  $A^\dagger$  or  $A$ . The energies of these new states will be  $(n+k+\frac{1}{2})\hbar\omega$  for some  $k \in \mathbb{Z}$ . However, we can't yet conclude that the energy levels are quantized, because we don't yet know that there isn't a continuum of possible starting-points  $|n\rangle$  (or indeed any!).

This brings us to the second key step: we must now investigate the norm of our states. Since  $N$  is Hermitian all its eigenvalues must certainly be real, but in fact they're also non-negative, because

$$n = n\langle n|n\rangle = \langle n|N|n\rangle = \langle n|A^\dagger A|n\rangle = \|A|n\rangle\|^2 \geq 0 \tag{3.11}$$

with  $n = 0$  iff  $A|n\rangle = 0$ , by properties of the norm. If  $A|n\rangle \neq 0$  then it is an eigenstate of  $N$  with eigenvalue  $n-1$ , but we've just shown that there are no states in  $\mathcal{H}$  with negative eigenvalue for  $N$ , so this lowering process must terminate. That will be the case iff  $n$  is a non-negative *integer*.

Putting all this together, we have a ground state  $|0\rangle$  of energy  $\frac{1}{2}\hbar\omega$  and an infinite tower of excited states  $|n\rangle$  of energy

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \quad \text{where } n \in \mathbb{N}_0. \tag{3.12}$$

These energy levels should be familiar from IB QM.

Acting on an energy eigenstate with the raising operator does not necessarily give us an excited state that is correctly normalised. In  $d = 1$ , energy eigenstates of the harmonic oscillator are non-degenerate<sup>16</sup>, so we must have  $A^\dagger|n\rangle = c_n|n+1\rangle$  for some constant  $c_n$  (which may depend on  $n$ ). Taking the norm of both sides shows that

$$|c_n|^2 = \|A^\dagger|n\rangle\|^2 = \langle n|AA^\dagger|n\rangle = \langle n|N+1|n\rangle = n+1. \quad (3.13)$$

Therefore, the correctly normalised state is

$$|n+1\rangle = \frac{1}{\sqrt{n+1}} A^\dagger|n\rangle = \frac{1}{\sqrt{(n+1)!}} (A^\dagger)^{n+1}|0\rangle. \quad (3.14)$$

Likewise, you can check that  $|n-1\rangle = \frac{1}{\sqrt{n}} A|n\rangle$  for  $n \geq 1$ , whilst again  $A|0\rangle = 0$ .

We've seen that everything about the energy levels of the harmonic oscillator follows from *i*) the algebra (*i.e.* the commutation relations) of raising and lowering operators, together with *ii*) considering the norm. Again, we'll follow these same two steps in analysing the eigenvalues and eigenstates of various operators throughout this course. This also parallels what you did in IB QM: there, by looking for series solutions of Schrödinger's equation, you could find an energy eigenstate for all  $E \in \mathbb{R}$ . However, these eigenstates were only normalizable if the series you found terminated. It was requiring this termination (*i.e.* normalizability) that lead to quantization of the energies.

I hope I've persuaded you that the algebraic approach above is somewhat cleaner and more efficient than looking for series solutions to Schrödinger's equation. Nothing has been lost, and we can easily use Dirac's formalism to recover the explicit wavefunctions of our eigenstates. For example, in the position representation, the defining equation  $A|0\rangle = 0$  of the ground state becomes

$$0 = \sqrt{2m\hbar\omega} \langle x|A|0\rangle = \langle x|m\omega X + iP|0\rangle = \left( m\omega x + \hbar \frac{d}{dx} \right) \psi_0(x), \quad (3.15)$$

---

<sup>16</sup>Here is a proof, which holds for any potential  $V(x)$  and energy  $E$  for which  $\exists x_0 > 0$  s.t.  $V(x) - E > 0 \forall |x| > x_0$ . In one dimension, the TISE  $-\hbar^2\psi''/2m + V(x)\psi = E\psi$  is a 2<sup>nd</sup>-order ode. Thus, for any fixed value of  $E$ , it has exactly two linearly independent solutions. Let  $\psi_1(x)$  and  $\psi_2(x)$  be the two solutions defined by the conditions

$$\begin{aligned} \psi_1(x_0) &= 1 & \text{and} & \psi'_1(x_0) = 0, \\ \psi_2(x_0) &= 0 & \text{and} & \psi'_2(x_0) = 1, \end{aligned}$$

where  $x_0$  is chosen so that  $V(x) - E \geq M^2 > 0$  for all  $|x| > x_0$  and some  $M > 0$ . Any solution of Schrödinger's equation for this  $V(x)$  and  $E$  is a linear combination of these two; our aim is to prove that (up to overall scale) there is only one combination that is normalizable.

Since they each solve the TISE, the Wronskian

$$W(\psi_1, \psi_2) = \psi_1\psi'_2 - \psi_2\psi'_1$$

is constant  $\forall x$ , and evaluating at  $x = x_0$  gives  $W(\psi_1, \psi_2) = 1$ . Also, our condition on  $V(x) - E$  implies that  $\psi_{1,2}(x)$  remain positive over the entire range  $x \in (x_0, \infty)$  and grow at least as fast as  $e^{Mx}$  as  $x \rightarrow \infty$ .

Since they are normalizable,  $\psi_{1,2}(x) \rightarrow 0$  as  $|x| \rightarrow \infty$ . By itself, this does not quite allow us to conclude that  $W = 0$ , since we need to know that  $\psi'$  are bounded. For this we'll need a condition on the potential, and it turns out that a sufficient criterion is that To see this, suppose

where  $\psi_0(x) = \langle x|0\rangle$  is the position space wavefunction of our ground state. This is a first order o.d.e. whose solution is

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}x^2\right), \quad (3.16)$$

where the overall constant is fixed (up to phase) by the normalization requirement  $\langle 0|0\rangle = \int_{\mathbb{R}} |\psi_0(x)|^2 = 1$ . This is the same Gaussian wavefunction for the ground state of the harmonic oscillator familiar from IB QM. If needed, the wavefunctions of the excited states may be found by acting with  $A^\dagger = (m\omega X - iP)/\sqrt{2m\hbar\omega}$ , which in the position representation is the differential operator  $\frac{1}{\sqrt{2}}(-\alpha\frac{d}{dx} + x/\alpha)$ , where  $\alpha = \sqrt{\hbar/m\omega}$ . You can check (or just Google) that these operators generate the usual Hermite polynomials.

### 3.2 Dynamics of Oscillators

At this stage, we've learned everything about the set  $\{|n\rangle\}$  of energy eigenstates of the quantum harmonic oscillator and their corresponding energies  $E_n = (n + \frac{1}{2})\hbar\omega$ . However, we still haven't said anything about the physics of how quantum oscillators actually *behave*. Classically, we know that a harmonic oscillator would undergo periodic motion with a period  $T = 2\pi/\omega$ . Furthermore, the energy of the classical oscillator is independent of the period, but is proportional to the square of the amplitude of oscillation. To what extent is the same true of our quantum oscillator?

To say anything about the *motion* of a quantum system, we need to examine the TDSE. To get started, first suppose our system is prepared at time  $t = 0$  to be in some energy eigenstate  $|n\rangle$ . Then by the TDSE, at a later time  $t$  it will have evolved to

$$|n, t\rangle = e^{-i(n+1/2)\omega t}|n\rangle. \quad (3.17)$$

Consequently, no eigenstate has a time dependence which oscillates at the classical frequency  $\omega$ . Even worse, no matter which energy eigenstate our oscillator is in, the expected position of the oscillator at any given time  $t$  is

$$\langle n, t | X | n, t \rangle = e^{+i(n+1/2)\omega t} \langle n | X | n \rangle e^{-i(n+1/2)\omega t} = \langle n | X | n \rangle, \quad (3.18)$$

where we've used the linearity / antilinearity properties of the inner product. The *rhs* is independent of time, so none of these states *move* – our oscillator does not appear to be oscillating!

To find some interesting dynamics, we must consider not a single energy eigenstate, but rather a *superposition*. This is a much more realistic assumption: there's no practical way we could prepare a macroscopic system to be in just one energy eigenstate. Let's now suppose our oscillator is prepared at  $t = 0$  to be in some generic state  $|\psi, 0\rangle = \sum_n c_n |n\rangle$ . The  $c_n$  should be chosen so that  $\langle \psi | \psi \rangle = 1$ , but are otherwise arbitrary. Then at time  $t$  this state will have evolved to

$$|\psi, t\rangle = \sum_{n=0}^{\infty} c_n e^{-iE_n t/\hbar} |n\rangle \quad (3.19)$$

by the TDSE.

Now let's examine where we expect to find such a generic state. We have

$$\begin{aligned}\langle \psi, t | X | \psi, t \rangle &= \left( \sum_m \bar{c}_m e^{iE_m t/\hbar} \langle m | \right) X \left( \sum_n c_n e^{-iE_n t/\hbar} | n \rangle \right) \\ &= \sum_{n,m} \bar{c}_m c_n e^{i(m-n)\omega t} \langle m | X | n \rangle.\end{aligned}\quad (3.20)$$

To evaluate the inner product  $\langle m | X | n \rangle$ , we write<sup>17</sup>

$$X = \sqrt{\frac{\hbar}{2m\omega}} (A + A^\dagger) \quad (3.21)$$

in terms of the raising and lowering operators. Recalling that  $A^\dagger | n \rangle = \sqrt{n+1} | n+1 \rangle$  and  $A | n \rangle = \sqrt{n} | n-1 \rangle$ , we have

$$\langle m | X | n \rangle = \sqrt{\frac{\hbar}{2m\omega}} (\sqrt{n} \langle m | n-1 \rangle + \sqrt{n+1} \langle m | n+1 \rangle), \quad (3.22)$$

showing that  $\langle m | X | n \rangle$  is non-zero only when  $m = n \pm 1$ . The double sum in (3.20) thus reduces to

$$\begin{aligned}\langle \psi, t | X | \psi, t \rangle &= \sqrt{\frac{\hbar}{2m\omega}} \left( \sum_{n=1}^{\infty} \sqrt{n} \bar{c}_{n-1} c_n e^{-i\omega t} + \sqrt{n} \bar{c}_n c_{n-1} e^{+i\omega t} \right) \\ &= \sum_n x_n \cos(\omega t + \phi_n),\end{aligned}\quad (3.23)$$

where the real numbers  $x_n$  and  $\phi_n$  are defined by

$$\sqrt{\frac{2n\hbar}{m\omega}} \bar{c}_{n-1} c_n = x_n e^{i\phi_n}. \quad (3.24)$$

Equation (3.23) shows that  $\langle X \rangle$  oscillates sinusoidally at exactly the classical frequency  $\omega$  whenever our oscillator is prepared in any generic<sup>18</sup> superposition of energy eigenstates. Furthermore, as for the classical oscillator, *the frequency of oscillation is independent of the energy*. In the calculation above, this occurs because the separation between every adjacent pair of energy levels is always  $\hbar\omega$ .

For a macroscopic oscillator, the only non-negligible amplitudes will be those where  $n \approx n_{\text{cl}}$  for some  $n_{\text{cl}} \gg 1$ . Consequently, a measurement of the energy is certain to yield some value close to  $E_{n_{\text{cl}}} = (n_{\text{cl}} + \frac{1}{2})\hbar\omega \approx n_{\text{cl}}\hbar\omega$ . For each eigenstate, we have

$$\langle n | X^2 | n \rangle = \frac{\hbar}{2m\omega} \langle n | AA^\dagger + A^\dagger A | n \rangle = \frac{E_n}{m\omega^2} \quad (3.25)$$

---

<sup>17</sup>To do this calculation using the techniques of IB QM, you would have worked in the position representation and said

$$\langle m | X | n \rangle = \int_{-\infty}^{\infty} \left( H_m(x) e^{-x^2/2\alpha} \right)^* x H_n(x) e^{-x^2/2\alpha} dx$$

where  $H_n(x)$  are Hermite polynomials of degree  $n$  and  $\alpha = \hbar/m\omega$ . This is correct, but the integral looks rather unpleasant to evaluate. Fortunately, our operator formalism means we never even have to try!

<sup>18</sup>Here, ‘generic’ simply means we must have non-zero amplitudes  $c_n = \langle n | \psi \rangle$  for at least one pair of adjacent energy levels.

so the mean value  $\overline{\langle X^2 \rangle} = E_{n_{\text{cl}}}/(m\omega^2)$ . Classically, the time average of  $x^2$  is proportional to the average potential energy, which for a harmonic oscillator is just half the total energy. Thus, classically we have  $\overline{x^2} = E/(m\omega^2)$  in agreement with the quantum result. The correspondence principle requires that the quantum and classical results agree for large values of  $E$ . That they agree even for low energies is a coincidence due to special symmetries of the harmonic oscillator.

### 3.2.1 Anharmonic Oscillations

Suppose that, instead of the pure harmonic oscillator, we have a potential that has a minimum at  $x = 0$  around which it grows approximately quadratically while  $|x| \ll a$ , but then asymptotes to a constant value at large  $|x|$ . Particles of low energy will have position space wavefunctions that are supported near the minimum of  $V(x)$ , so we'd expect the low-lying energy levels to be roughly equal to those of a corresponding harmonic oscillator. Particles of higher energy would start to see the fact that the potential is not purely harmonic. In fact, for any such asymptotically constant potential with a single extremum, the separation between energy levels gets smaller and smaller as we approach the asymptotic value  $\lim_{x \rightarrow \infty} V(x)$  of the potential (beyond which we have a continuum of non-bound states)<sup>19</sup>.

Let's prepare a state to be in two adjacent energy levels, say

$$|\psi\rangle = c_n|n\rangle + c_{n+1}|n+1\rangle,$$

where  $|n\rangle$  is the  $n^{\text{th}}$  energy eigenstate of our anharmonic potential, whatever it may be. Since the potential is symmetric around  $x = 0$ , its eigenstates have definite parity and so  $\langle n|X|n\rangle = 0$  for all  $|n\rangle$ . Therefore, at time  $t$ ,

$$\langle \psi, t | X | \psi, t \rangle = \bar{c}_n c_{n+1} e^{i(E_n - E_{n+1})t/\hbar} \langle n | X | n+1 \rangle + \text{c.c.} \quad (3.26)$$

This is again a sinusoidally oscillating function of time, but now the period  $T = 2\pi\hbar/(E_{n+1} - E_n)$  depends on  $n$  since we no longer expect the energy levels to be equally spaced. In fact, since the energy levels get closer together as  $n$  increases, if we give our particle a larger amplitude of oscillation – and hence more energy – it will take longer to execute a complete an oscillation. This is just what we'd expect classically.

For example, consider the *Pöschl-Teller* potential

$$V(X) = -V_0 \operatorname{sech}^2(\kappa X) \quad (3.27)$$

for some constant  $V_0 > 0$  and length scale  $a = 1/\kappa$  (see figure 3). This has bound state energy levels<sup>20</sup>

$$E_n = -\frac{\hbar^2 \kappa^2}{2m} (\nu - n)^2 \quad \text{for } 0 \leq n < \nu, \quad (3.28)$$

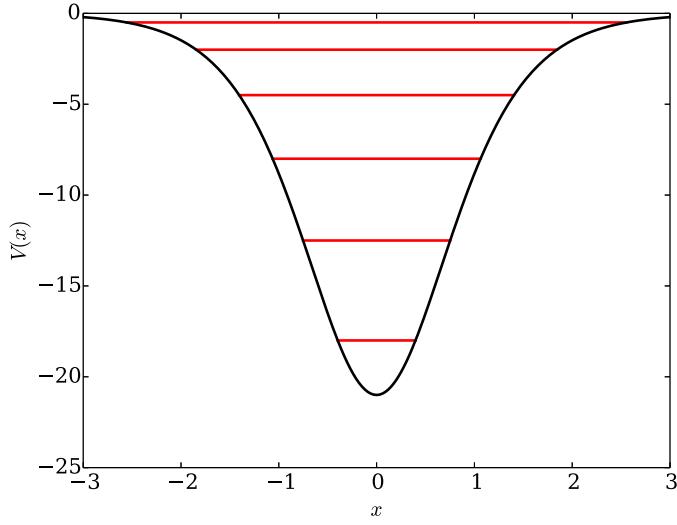
where  $\nu$  is the positive root of  $\nu(\nu + 1) = 2mV_0/\hbar^2 \kappa^2$ . The separation between adjacent bound state energy levels is thus

$$E_{n+1} - E_n = (2(\nu - n) - 1)\hbar^2 \kappa^2 / 2m \quad (3.29)$$

---

<sup>19</sup>For a proof, see *e.g.* chapter 6 of S. Gustafson & I.M. Singer, *Mathematical Concepts of Quantum Mechanics*, Springer (2010). (Both this result and its proof are non-examinable in this course.)

<sup>20</sup>Exercise: Try to prove this. It's similar to one of the questions on [Problem Sheet 1](#).



**Figure 3:** The Pöschl-Teller potential  $V(x) = -V_0 \operatorname{sech}^2(\kappa x)$  and its energy levels. Figure by Nicoguaro, taken from [this](#) Wikipedia page.

and decreases as  $n$  increases towards  $\nu$ . (These formulae break down when  $n \geq \nu$ , where we enter the continuum of non-normalizable states.)

If  $V_0 \gg \hbar^2 \kappa^2 / 2m$  then the potential is very deep and contains many bound states. In this regime, we have  $\nu \approx \sqrt{2mV_0/\hbar^2\kappa^2} \gg 1$  and so from (3.29) we find that a superposition of low-lying states will oscillate with a frequency

$$\omega \approx \frac{\hbar\kappa^2}{m}\nu \approx \kappa\sqrt{\frac{2V_0}{m}}. \quad (3.30)$$

For  $x \ll 1/\kappa$ , we may approximate  $-V_0 \operatorname{sech}^2(\kappa x) \approx -V_0 + V_0 \kappa^2 x^2$ , so this frequency is indeed just what we'd expect for the corresponding harmonic oscillator. On the other hand, a superposition of states clustered around  $n = \nu/2$  will oscillate at around half this frequency as neighbouring energy levels in this region are only separated by about half as much.

If we include a wider range of states in our initial superposition, then we'll instead find a sequence of terms<sup>21</sup>

$$\begin{aligned} \langle \psi, t | X | \psi, t \rangle &= \cdots + \bar{c}_{n-1} c_n e^{i(E_{n-1} - E_n)t/\hbar} \langle n-1 | X | n \rangle + \bar{c}_{n+1} c_n e^{i(E_{n+1} - E_n)t/\hbar} \langle n+1 | X | n \rangle \\ &\quad + \bar{c}_{n+3} c_n e^{i(E_{n+3} - E_n)t/\hbar} \langle n+3 | X | n \rangle + \cdots. \end{aligned} \quad (3.31)$$

Since this is no longer a harmonic oscillator, we do not generically expect  $\langle n+3 | X | n \rangle = 0$ . Let's define a frequency  $\Omega = (E_{n_{\text{cl}}+1} - E_{n_{\text{cl}}})/\hbar$  where  $n_{\text{cl}} \gg 1$ . Provided the  $c_n$ 's are clustered around this large central value  $n = n_{\text{cl}}$  sufficiently tightly that the difference between adjacent energy levels is roughly constant over the range of  $n$  for which the  $c_n$  are

---

<sup>21</sup>Note that for a symmetric potential, we again have  $\langle n+2k | X | n \rangle = 0$  by parity.

appreciable, then, to reasonable accuracy, all the terms that contribute to (3.31) oscillate with frequencies that are integer multiples of  $\Omega_{n_{\text{cl}}}$ . Thus the motion will be periodic, but anharmonic, just as we expect classically.

If we release the anharmonic oscillator from some large extension  $x_0$ , then initially the wavefunction will be a superposition of many energy levels, with coefficients that ensure  $\langle x|\psi\rangle = \sum_n c_n \langle x|n\rangle$  is sharply peaked around  $x_0$ . At time  $t$ , this state will evolve to

$$|\psi, t\rangle = e^{-iE_{n_{\text{cl}}}t/\hbar} \sum_n c_n e^{i(E_{n_{\text{cl}}}-E_n)t/\hbar} |n\rangle.$$

Since the separation between energy levels varies, the frequencies appearing in this sum are not all integer multiples of any  $\Omega_N$ . Consequently, after a time of order  $2\pi/\Omega_N$ , most terms in the sum will have not quite returned to their original values, so the wavefunction at  $t = 2\pi/\Omega$  will be less sharply peaked around  $X$ . With each subsequent passing of time  $2\pi/\Omega_N$ , the wavefunction will become more and more diffusely spread. Classically, if we release an oscillator with a rather uncertain value for it's energy, then for a pure harmonic oscillator we always know where to find the oscillator at later times, since it's period is independent of the energy. However, for an anharmonic oscillator, the period depends on the energy, so after a long time our oscillator is equally likely to be located anywhere.

## 4 Transformations and Symmetries

While classical physics is firmly rooted in space-time, quantum mechanics takes place in the more abstract Hilbert space. Actions such as translations or rotations around a given origin have a straightforward space-time interpretation, but how can these affect a quantum particle living in Hilbert space?

In this chapter, we'll return to the general formalism of QM. We'll understand how transformations such as spatial translations or rotations affect states in Hilbert space, linking together the world of quantum mechanics with our familiar experience in ‘normal’ space  $\mathbb{R}^3$ . In so doing, we'll obtain a deeper appreciation of the origin of many of the most common commutation relations, including most of the ones you're familiar with from 1B. We'll also understand *why* the momentum operator  $\hat{P} = -i\hbar \partial/\partial x$ , a fact which probably seemed rather mysterious in 1B QM.

### 4.1 Transformations of States and Operators

If we apply a transformation such as a translation or rotation to our physical system, we'd expect states to be different before and after the transformation. For example, if a particle is prepared in a state  $|\psi\rangle$  that is strongly peaked around the origin  $\mathbf{0} \in \mathbb{R}^3$  and we translate our system through a vector  $\mathbf{a}$ , after this transformation we should expect to find our particle in a new state  $|\psi'\rangle$  whose wavefunction is strongly peaked around  $\mathbf{x} = \mathbf{a}$ . This suggests that the effect of any given spatial transformation should be represented on Hilbert space by a linear operator

$$U : \mathcal{H} \rightarrow \mathcal{H} \quad U : |\psi\rangle \mapsto |\psi'\rangle = U|\psi\rangle. \quad (4.1)$$

where the operator in question depends on the specific type of transformation.

Whatever state we start with, after applying the transformation, we still expect to find it *somewhere* in space. Thus, properly normalised states should remain so after applying  $U$ , or in other words

$$1 = \langle\psi|\psi\rangle = \langle\psi'|\psi'\rangle = \langle\psi|U^\dagger U|\psi\rangle \quad \text{for all } |\psi\rangle \in \mathcal{H}. \quad (4.2)$$

We'll now show that the requirement that  $U$  preserves the norm of our state fixes it to be a *unitary* operator<sup>16</sup>. The required trick is sometimes called the *polarization identity*. We first write  $|\psi\rangle$  as  $|\psi\rangle = |\phi\rangle + \lambda|\chi\rangle$  for some states  $|\phi\rangle, |\chi\rangle \in \mathcal{H}$  and some  $\lambda \in \mathbb{C}$ . Equation (4.2) becomes

$$\langle\phi|\phi\rangle + \bar{\lambda}\langle\chi|\phi\rangle + \lambda\langle\phi|\chi\rangle + |\lambda|^2\langle\chi|\chi\rangle = \langle\phi|U^\dagger U|\phi\rangle + \bar{\lambda}\langle\chi|U^\dagger U|\phi\rangle + \lambda\langle\phi|U^\dagger U|\chi\rangle + |\lambda|^2\langle\chi|U^\dagger U|\chi\rangle \quad (4.3)$$

---

<sup>16</sup>Because physical states are represented by rays in  $\mathcal{H}$ , rather than actual vectors, there is one other possibility, transformations may be represented by operators that are both *anti-linear*, in the sense that  $U(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = \bar{c}_1 U|\psi_1\rangle + \bar{c}_2 U|\psi_2\rangle$ , and *anti-unitary*, meaning that  $\langle\chi|U^\dagger U|\phi\rangle = \langle\phi|\chi\rangle = \overline{\langle\chi|\phi\rangle}$  for all  $|\phi\rangle, |\chi\rangle \in \mathcal{H}$ . Such antiunitary operators turn out to be related to reversing the direction of time; we will not discuss them further in this course. Wigner showed that all transformations are represented by either linear, unitary operators or anti-linear, anti-unitary ones.

where we recall that the inner product is antilinear in its left entry. By assumption  $\langle \phi | U^\dagger U | \phi \rangle = \langle \phi | \phi \rangle$  and  $\langle \chi | U^\dagger U | \chi \rangle = \langle \chi | \chi \rangle$ , so this simplifies to

$$\lambda (\langle \phi | \chi \rangle - \langle \phi | U^\dagger U | \chi \rangle) = \bar{\lambda} (\langle \chi | U^\dagger U | \phi \rangle - \langle \chi | \phi \rangle). \quad (4.4)$$

In order for this to hold for *every*  $|\psi\rangle \in \mathcal{H}$ , it must continue to hold as we vary  $\lambda$ . In particular, as the phase of  $\lambda$  varies, the only way for (4.4) to be satisfied is if

$$\langle \phi | \chi \rangle = \langle \phi | U^\dagger U | \chi \rangle. \quad (4.5)$$

Finally, for this hold for arbitrary states  $|\phi\rangle$  and  $|\chi\rangle$  we must have  $U^\dagger U = 1$ . Multiplying through on the right by  $U^{-1}$  shows that

$$U^\dagger = U^{-1} \quad (4.6)$$

which says that the adjoint of  $U$  is equal to its inverse. Operators with this property are said to be *unitary*<sup>17</sup>.

There's one further condition on our operators  $U$ . Transformations of space such as translations and rotations form a *group*. For example, the composition of two rotations is again a rotation, the trivial rotation forms the identity, and the inverse of a rotation is a rotation of the same amount around the same axis, but in the opposite sense (*e.g.* clockwise instead of anticlockwise). Let's suppose that our spatial transformations form a group  $G$ . To reflect this group structure, our operators should<sup>18</sup> provide a *homomorphism* from  $G$  to the group of unitary operators in the sense that, for all  $g_1, g_2 \in G$ ,

$$U(g_2) \circ U(g_1) = U(g_2 \cdot g_1) \quad \text{and} \quad U(1) = 1_{\mathcal{H}} \quad (4.7)$$

where  $\cdot$  denotes the group multiplication in  $G$  and  $\circ$  denotes the composition of linear operators acting on  $\mathcal{H}$ . (We'll typically suppress these composition symbols from now on.) Note that if  $U_1$  and  $U_2$  are each unitary operators, then  $(U_2 U_1)^{-1} = U_1^{-1} U_2^{-1} = U_1^\dagger U_2^\dagger = (U_2 U_1)^\dagger$  and so the composite operator  $U_2 U_1$  is also unitary. Note also that the identity operator  $1_{\mathcal{H}}$  is trivially unitary.

A particularly important class of transformations are those that depend smoothly on some parameter  $\theta$ . For example, we can smoothly vary both the axis about which and the angle through which we rotate, or the magnitude and direction of a translation vector. If  $\theta = 0$  is trivial transformation, represented on  $\mathcal{H}$  by the identity operator, then for infinitesimal transformation we have

$$U(\delta\theta) = 1 - i\delta\theta T + O(\delta\theta^2) \quad (4.8)$$

where  $T$  is some operator that is independent of  $\theta$ . (The factor of  $-i$  in this equation is just a convention for later convenience.)  $T$  is called the *generator* of the transformation. For such infinitesimal transformations, the condition (4.6) that  $U$  is unitary becomes

$$1 + i\delta\theta T^\dagger + O(\delta\theta^2) = 1 + i\delta\theta T + O(\delta\theta^2), \quad (4.9)$$

---

<sup>17</sup>Note that unitary operators are certainly bounded, and in fact have unit norm in the operator topology.

<sup>18</sup>This statement is not quite accurate – there's an important refinement that we'll return to later in the course.

which to first order in  $\delta\theta$  gives

$$T = T^\dagger. \quad (4.10)$$

Thus the generator  $T$  is Hermitian and hence a good candidate for an observable quantity. A finite transformation can be generated by repeatedly performing an infinitesimal one. Specifically, if we set  $\delta\theta = \theta/N$  and transform  $N$  times with  $U(\delta\theta)$ , then in the limit  $N \rightarrow \infty$  we have

$$U(\theta) = \lim_{N \rightarrow \infty} \left( 1 - i \frac{\theta}{N} T \right)^N = e^{-i\theta T}, \quad (4.11)$$

where the exponential of an operator may be defined by (2.42), or equivalently by its power series expansion. This form of the unitary operator  $U(\theta)$  is especially useful when states are expressed in terms of the basis of eigenstates of the Hermitian generator  $T$ .

We obtain an important equation by using (4.8) to evaluate  $|\psi'\rangle = U(\delta\theta)|\psi\rangle$ . Subtracting  $|\psi\rangle$  from both sides and dividing through by  $\delta\theta$ , in the limit  $\delta\theta \rightarrow 0$  we obtain

$$i \frac{\partial |\psi\rangle}{\partial \theta} = T|\psi\rangle \quad (4.12)$$

There is no assumption here that the wavefunction  $\psi(\mathbf{x}) = \langle \mathbf{x} | \psi \rangle$  should be differentiable as a function of space. We merely say that as our transformation varies smoothly away from the identity, so too does the state  $|\psi\rangle$  vary smoothly inside  $\mathcal{H}$ . The rate at which  $|\psi\rangle$  varies is governed by the generator.

Having described how transformations act on states in  $\mathcal{H}$ , we should now understand how they act on operators. Suppose  $A$  is some operator whose matrix elements we are interested in. If our transformation maps  $|\psi\rangle \rightarrow |\psi'\rangle$ , then the expectation value of  $A$  will be mapped as

$$\langle \psi | A | \psi \rangle \mapsto \langle \psi' | A | \psi' \rangle = \langle \psi | U^\dagger(\theta) A U(\theta) | \psi \rangle. \quad (4.13)$$

Consequently, we can find the expectation value  $A$  will take after the transformation by working with the original states, but instead transforming the operator as

$$A \mapsto A' := U^\dagger(\theta) A U(\theta) = U^{-1}(\theta) A U(\theta), \quad (4.14)$$

using the fact that  $U$  is unitary. This is known as a *similarity transform*. Note that

$$[A', B'] = [U^{-1} A U, U^{-1} B U] = U^{-1} [A, B] U, \quad (4.15)$$

so the commutator of two similarity transforms is the similarity transform of the commutator. Also, similarity transforms do not change the spectrum of any operator: if  $|\psi\rangle$  is an eigenstate of  $A$  with eigenvalue  $a$ , then  $U^{-1}|\psi\rangle$  is an eigenstate of  $A'$  with the same eigenvalue.

Finally, for an infinitesimal transformation, we have

$$U^\dagger(\delta\theta) A U(\delta\theta) = A + i \delta\theta [T, A] + O(\delta\theta^2), \quad (4.16)$$

so  $\delta A = i \delta\theta [T, A]$ . Thus, while the rate of change of states themselves is given by the action of the generator, the rate of change of operators under a transformation is determined by

the *commutator* of the generator with the operator. We'll see that most of the commutation relations we meet in quantum mechanics (and certainly all the familiar ones) arise this way. Furthermore, because Hermitian operators represent observable quantities with which we are familiar, in practice it's often easier to understand how a transformation should act on an operator, rather than on the more abstract notion of a state in Hilbert space.

The above discussion has been rather abstract. Let's ground it by looking at a few of the most important examples of transformations and their corresponding generators.

## 4.2 Translations

When we move an object around, we expect to find it in a new place. Specifically, suppose  $\langle \psi | \mathbf{X} | \psi \rangle = \mathbf{x}_0$  for some normalised state  $|\psi\rangle$ . Since  $\mathbf{x}_0$  just labels a spatial point, it must behave under translations and rotations like any vector. Thus, after the translation we expect our object to be described by a new state  $|\psi'\rangle$  with  $\langle \psi' | \mathbf{X} | \psi' \rangle = \mathbf{x}_0 + \mathbf{a}$ . The general theory of the previous section asserts that this translation is represented on Hilbert space by a unitary operator<sup>19</sup>  $U(\mathbf{a})$  so that  $|\psi'\rangle = U(\mathbf{a})|\psi\rangle$ . Therefore we can write the expected location of the translated state as

$$\begin{aligned}\langle \psi' | \mathbf{X} | \psi' \rangle &= \langle \psi | \mathbf{X} | \psi \rangle + \mathbf{a} = \langle \psi | (\mathbf{X} + 1_{\mathcal{H}} \mathbf{a}) | \psi \rangle \\ &= \langle \psi | U^{-1}(\mathbf{a}) \mathbf{X} U(\mathbf{a}) | \psi \rangle,\end{aligned}\tag{4.17}$$

where  $1_{\mathcal{H}}$  represents the identity operator on  $\mathcal{H}$ . (We'll often drop the identity operator where it's unambiguous.) Since this is true for any state  $|\psi\rangle$ , the same argument as above shows that

$$U^{-1}(\mathbf{a}) \mathbf{X} U(\mathbf{a}) = \mathbf{X} + \mathbf{a} 1_{\mathcal{H}}.\tag{4.18}$$

Let's be clear about what this equation means. In this course, we're interested in Quantum Mechanics on  $\mathbb{R}^3$ , so the (boldface) position operator  $\mathbf{X} : \mathcal{H} \rightarrow \mathcal{H}$  is really a collection of three operators corresponding to the three coordinates of  $\mathbb{R}^3$ . Taking the standard Cartesian basis of  $\mathbb{R}^3$ , in more detail (4.18) says

$$\begin{pmatrix} U^{-1}(\mathbf{a}) X U(\mathbf{a}) \\ U^{-1}(\mathbf{a}) Y U(\mathbf{a}) \\ U^{-1}(\mathbf{a}) Z U(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} X + a_x 1_{\mathcal{H}} \\ Y + a_y 1_{\mathcal{H}} \\ Z + a_z 1_{\mathcal{H}} \end{pmatrix}.\tag{4.19}$$

In other words, conjugating the  $X$  (component!) operator by  $U(\mathbf{a})$  gives a new operator  $U^{-1}(\mathbf{a}) X U(\mathbf{a}) : \mathcal{H} \rightarrow \mathcal{H}$  that we identify with  $X + a_x$ , and similarly for  $Y$  and  $Z$ . If  $U(\mathbf{a})$  acted on the position operator in any other way, it could not represent a translation, so equation (4.18) may be taken as the defining property of the translation operator, distinguishing it from other unitary transformations.

---

<sup>19</sup>Strictly, we should label this as  $U_{\text{trans}}(\mathbf{a})$  to emphasise that it is the unitary operator generating *translations*, as opposed to anything else we might associate to  $\mathbf{a}$ . We typically drop such additional labels. Exactly which unitary operator we're talking about should hopefully be clear from the context.

For an infinitesimal translation  $\delta\mathbf{a}$ , equation (4.8) states that the translation operator can be written as<sup>20</sup>

$$U(\delta\mathbf{a}) = 1 - \frac{i}{\hbar} \delta\mathbf{a} \cdot \mathbf{P} + O(\delta\mathbf{a}^2) \quad (4.20)$$

where  $\mathbf{P}$  is Hermitian, I've named this operator  $\mathbf{P}$ , so I'm not going to be able to fool any of you that it will turn out to be anything other than the momentum operator. But that's not how I want you to think of it yet. *By definition*,  $\mathbf{P}/\hbar$  is the generator of infinitesimal translations.

The translation generator  $\mathbf{P}/\hbar$  must have units  $1/(\text{length})$  so that it makes sense to add the second term of (4.20) to the first. Because we've included a factor of  $\hbar$ , our  $\mathbf{P}$  must have dimensions of  $(\text{mass} \times \text{velocity})$ . In fact, this is just a convention. We could choose to measure masses in units of  $1/(\text{length} \times \text{velocity})$ , using  $\hbar$  as a conversion factor for our units. In relativity, the speed of light  $c$  provides a further natural conversion factor between lengths and times, so in units of  $\hbar/c$ , masses are equivalent to inverse lengths. In *natural units*, we work with length and time scales adapted to Nature rather than our civilisation, so we set  $\hbar = 1$  and  $c = 1$ . In these units, the translation generator is precisely  $\mathbf{P}$ . Incidentally, you may be worried about setting  $\hbar = 1$  when it is such a ‘small’ number;  $\hbar \approx 1.0547 \times 10^{-34} \text{ Js/rad}$ . But this is quite wrong. Every atom in the Universe knows that  $\hbar$  is not at all small or insignificant. It only appears small when measured in units such as Joules and seconds that are relevant for steam engines and pocketwatches. The real, deep question is not why  $\hbar$  appears small, but why humans are so big.

Plugging (4.20) into the defining equation (4.18) for the translation operator we find

$$\frac{i}{\hbar} [\delta\mathbf{a} \cdot \mathbf{P}, \mathbf{X}] = \delta\mathbf{a} \quad (4.21)$$

and since this holds for any infinitesimal translation,

$$[X_i, P_j] = i\hbar \delta_{ij}. \quad (4.22)$$

These, as I'm sure you recognise, are the fundamental commutation relations of quantum mechanics. I'm prepared to bet that the first time you met them (and perhaps even up until now) you thought they were a weird and mysterious feature of quantum mechanics. Here they're revealed as little more than common sense: asking where you are and then translating is not the same as first translating and then asking where you are. What *is* weird and quantum about this equation is *not* that  $\mathbf{X}$  and  $\mathbf{P}$  don't commute, but the fact that they are operators acting on a Hilbert space.

By repeatedly performing the same infinitesimal translation, we can write

$$U(\mathbf{a}) = \exp\left(-\frac{i}{\hbar}\mathbf{a} \cdot \mathbf{P}\right) \quad (4.23)$$

for a translation through the finite vector  $\mathbf{a}$ . Since  $U(\mathbf{a})$  provides a homomorphism of this group of translations to the group of linear operators on  $\mathcal{H}$ , we have

$$U(\mathbf{b}) U(\mathbf{a}) = U(\mathbf{b} + \mathbf{a}) = U(\mathbf{a} + \mathbf{b}) = U(\mathbf{a}) U(\mathbf{b}), \quad (4.24)$$

---

<sup>20</sup>Here,  $\delta\mathbf{a} \cdot \mathbf{P} = \sum_{i=1}^3 \delta a_i P_i$  involves the standard dot product in  $\mathbb{R}^3$ . Note that the result is still an operator on  $\mathcal{H}$ , a linear combination of the three operators  $\mathbf{P}$ .

where in the second equality we used the fact that  $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ , stating that the order in which we perform two translations doesn't matter<sup>21</sup>. Using (4.20) to expand the far left & far right of this equation to lowest non-trivial order in  $a_i b_j$  shows that the generators  $\mathbf{P}$  obey

$$[P_i, P_j] = 0 \quad \text{for all } i, j \quad (4.25)$$

and so commute with each other. Again, the fundamental meaning of this equation is simply that translations form an Abelian group — the order of translations doesn't matter. It only says anything about our ability to simultaneously know all three components of a particle's momentum once we identify  $\mathbf{P}$  as corresponding to momentum, which we haven't yet done.

Now let's think about what the translation operator does to states. Let  $|\mathbf{x}\rangle$  represent an eigenstate of the position operator with eigenvalue  $\mathbf{x}$ , so that  $\mathbf{X}|\mathbf{x}\rangle = \mathbf{x}|\mathbf{x}\rangle$ , with normalisation condition

$$\langle \mathbf{x}' | \mathbf{x} \rangle = \delta^3(\mathbf{x}' - \mathbf{x}) \quad (4.26)$$

as usual for continuum states<sup>22</sup>. This state represents a particle which is definitely located at  $\mathbf{x} \in \mathbb{R}^3$ . Applying the translation operator,

$$\mathbf{X}(U(\mathbf{a})|\mathbf{x}\rangle) = ([\mathbf{X}, U(\mathbf{a})] + U(\mathbf{a})\mathbf{X})|\mathbf{x}\rangle. \quad (4.27)$$

We can evaluate the commutator by multiplying (4.18) through by  $U(\mathbf{a})$  on the left, finding  $[\mathbf{X}, U(\mathbf{a})] = U(\mathbf{a})\mathbf{a}$ . Since  $\mathbf{a}$  is just a constant vector (not an operator on  $\mathcal{H}$ ) it commutes with the translation operator, so

$$\mathbf{X}(U(\mathbf{a})|\mathbf{x}\rangle) = (\mathbf{x} + \mathbf{a})(U(\mathbf{a})|\mathbf{x}\rangle). \quad (4.28)$$

As anticipated, our new state  $U(\mathbf{a})|\mathbf{x}\rangle$  is certainly located at  $\mathbf{x} + \mathbf{a}$ . Consequently,  $U(\mathbf{a})|\mathbf{x}\rangle$  must be proportional to the state  $|\mathbf{x} + \mathbf{a}\rangle$ , the normalised state that is definitely located at  $\mathbf{x} + \mathbf{a}$ . Setting  $U(\mathbf{a})|\mathbf{x}\rangle = c|\mathbf{x} + \mathbf{a}\rangle$  for some  $c \in \mathbb{C}$  and taking the inner product with another position eigenstate  $|\mathbf{x}'\rangle$  we have<sup>23</sup>

$$\begin{aligned} c \delta^3(\mathbf{x}' - \mathbf{x} - \mathbf{a}) &= c \langle \mathbf{x}' | \mathbf{x} + \mathbf{a} \rangle = \langle \mathbf{x}' | U(\mathbf{a}) | \mathbf{x} \rangle = (U^{-1}(\mathbf{a}) | \mathbf{x}' \rangle)^\dagger | \mathbf{x} \rangle \\ &= \left( \frac{1}{c} | \mathbf{x}' - \mathbf{a} \rangle \right)^\dagger | \mathbf{x} \rangle = \frac{1}{c} \delta^3(\mathbf{x}' - \mathbf{a} - \mathbf{x}) \end{aligned} \quad (4.29)$$

---

<sup>21</sup>Here this is of course the statement that the group of spatial translations is  $(\mathbb{R}^3, +)$ , where the group operation '+' is commutative. Groups for which the group operation is commutative are often called *Abelian*.

<sup>22</sup>Technically,  $|\mathbf{x}\rangle$  is a non-normalizable state, which really means it doesn't live in the Hilbert space  $L^2(\mathbb{R}^3, d^3x)$ . This is related to the fact that the position operator  $\mathbf{X}$  is *unbounded* — there is no constant  $M$  such that  $\|\mathbf{X}|\psi\rangle\| \leq M\||\psi\rangle\|$  for all  $|\psi\rangle \in \mathcal{H}$ . Naively, this means that ‘the eigenvalues of  $\mathbf{X}$  can be arbitrarily large’, which is physically reasonable since our particle may be located very far away. However, in general, unbounded operators do not have any eigenstates in  $\mathcal{H}$ , so do not really have eigenvalues. The distinction between bounded and unbounded operators is tremendously important in functional analysis, but will not play a role in this course.

<sup>23</sup>Note the slight awkwardness of the Dirac notation here. Because operators act to the right, we have to write  $\langle \chi | U | \phi \rangle = (U^{-1} | \chi \rangle)^\dagger | \phi \rangle$  for any unitary operator  $U$ , as we used in the third equality. In mathematical notation we would simply write  $(\chi, U\phi) = (U^{-1}\chi, \phi)$ .

where we used the fact that  $U(\mathbf{a})$  is unitary and the fact that if  $U(\mathbf{a})|\mathbf{x}\rangle = c|\mathbf{x} + \mathbf{a}\rangle$  then  $U^{-1}(\mathbf{a})|\mathbf{x} + \mathbf{a}\rangle = c^{-1}|\mathbf{x}\rangle$ . Comparing both sides shows that  $|c|^2 = 1$ , so  $c$  is a pure phase and without loss of generality we can set  $c = 1$ .

Now suppose we consider translating an arbitrary state  $|\psi\rangle$ . The position space wavefunction of this state is simply its coefficient

$$\psi(\mathbf{x}) = \langle \mathbf{x} | \psi \rangle \quad (4.30)$$

in the position basis. Thus the wavefunction  $\psi_{\text{trans}}(\mathbf{x})$  of the translated state  $|\psi_{\text{trans}}\rangle = U(\mathbf{a})|\psi\rangle$  is

$$\begin{aligned} \psi_{\text{trans}}(\mathbf{x}) &= \langle \mathbf{x} | \psi_{\text{trans}} \rangle = \langle \mathbf{x} | U(\mathbf{a}) | \psi \rangle = (U^{-1}(\mathbf{a}) | \mathbf{x} \rangle)^\dagger | \psi \rangle \\ &= \langle \mathbf{x} - \mathbf{a} | \psi \rangle = \psi(\mathbf{x} - \mathbf{a}). \end{aligned} \quad (4.31)$$

Eq. (4.31) says that the wavefunction of the translated state takes the same value at  $\mathbf{x}$  as the original wavefunction took at  $\mathbf{x} - \mathbf{a}$ . This is completely natural as we've translated our state through  $\mathbf{a}$ !

In particular, for an infinitesimal translation  $\delta\mathbf{a}$ , on the one hand we can Taylor expand the translated wavefunction to find<sup>24</sup>

$$\psi_{\text{trans}}(\mathbf{x}) - \psi(\mathbf{x}) = -\delta\mathbf{a} \cdot \nabla \psi(\mathbf{x}), \quad (4.32)$$

while on the other we expand the operator  $U(\delta\mathbf{a})$  to find

$$\psi_{\text{trans}}(\mathbf{x}) - \psi(\mathbf{x}) = \left\langle \mathbf{x} \left| 1 - \frac{i}{\hbar} \delta\mathbf{a} \cdot \mathbf{P} \right| \psi \right\rangle - \langle \mathbf{x} | \psi \rangle = -\frac{i}{\hbar} \delta\mathbf{a} \cdot \langle \mathbf{x} | \mathbf{P} | \psi \rangle \quad (4.33)$$

using (4.20) to lowest order in  $\delta\mathbf{a}$ . Consequently, for any state  $|\psi\rangle$  the momentum operator  $\mathbf{P}$  acts in the position representation as

$$\langle \mathbf{x} | \mathbf{P} | \psi \rangle = -i\hbar \nabla \psi(\mathbf{x}). \quad (4.34)$$

This is just what you would have said in IB QM last year, here derived from the point of view of the effect a translation of  $\mathbb{R}^3$  has on states in Hilbert space.

As a special case, let's apply this argument to the state  $|\mathbf{p}\rangle$  that obeys<sup>25</sup>  $\mathbf{P}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle$ , representing a particle whose momentum is certainly  $\mathbf{p}$ . In this case the above argument becomes

$$\begin{aligned} \psi_{\mathbf{p}}(\mathbf{x} - \mathbf{a}) &= \langle \mathbf{x} - \mathbf{a} | \mathbf{p} \rangle = \langle \mathbf{x} | U(\mathbf{a}) | \mathbf{p} \rangle = \langle \mathbf{x} | e^{-i\mathbf{a} \cdot \mathbf{P}/\hbar} | \mathbf{p} \rangle \\ &= e^{-i\mathbf{a} \cdot \mathbf{p}/\hbar} \langle \mathbf{x} | \mathbf{p} \rangle = e^{-i\mathbf{a} \cdot \mathbf{p}/\hbar} \psi_{\mathbf{p}}(\mathbf{x}), \end{aligned} \quad (4.35)$$

where in going to the second line we used the fact that  $|\mathbf{p}\rangle$  is an eigenstate of  $\mathbf{P}$ . Comparing the first and last expression, we deduce that the position space wavefunction of a momentum eigenstate must take the form

$$\psi_{\mathbf{p}}(\mathbf{x}) = C e^{i\mathbf{p} \cdot \mathbf{x}/\hbar} \quad (4.36)$$

---

<sup>24</sup>At least if the wavefunction is sufficiently smooth.

<sup>25</sup>As with the position operator, the momentum operator  $\mathbf{P}$  is unbounded, so the ‘momentum eigenstates’  $|\mathbf{p}\rangle$  are necessarily non-normalizable, as we'll soon see explicitly.



**Figure 4:** An infinitesimal rotation transforms  $\mathbf{v} \mapsto \mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v}$ .

for some  $\mathbf{x}$ -independent factor  $C$ . It's convenient to choose  $C$  to be  $(2\pi\hbar)^{-3/2}$  since in this case we have

$$\begin{aligned}\langle \mathbf{p}' | \mathbf{p} \rangle &= \int d^3y d^3x \langle \mathbf{p}' | \mathbf{y} \rangle \langle \mathbf{y} | \mathbf{x} \rangle \langle \mathbf{x} | \mathbf{p} \rangle = \int \frac{d^3x d^3y}{(2\pi\hbar)^3} e^{-i\mathbf{p}' \cdot \mathbf{y}/\hbar} \delta^3(\mathbf{y} - \mathbf{x}) e^{i\mathbf{p} \cdot \mathbf{x}/\hbar} \\ &= \int \frac{d^3x}{(2\pi\hbar)^3} e^{-i(\mathbf{p}' - \mathbf{p}) \cdot \mathbf{x}/\hbar} = \delta^3(\mathbf{p}' - \mathbf{p})\end{aligned}\quad (4.37)$$

so that the momentum eigenstates are normalised in the same way as the position eigenstates (and in the same way as for any continuum states). Again, the form

$$\psi_{\mathbf{p}}(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p} \cdot \mathbf{x}/\hbar} \quad (4.38)$$

of the wavefunction for momentum eigenstates is familiar, but here we've derived it from first principles rather than using the position space representation (4.34) of the momentum operator. (We assumed this result earlier when showing that the position and momentum space wavefunctions are each other's Fourier transforms.) Note also that  $|\psi_{\mathbf{p}}(\mathbf{x})|^2 = 1/(2\pi\hbar)^3$ , so  $\int_{\mathbb{R}^3} |\psi_{\mathbf{p}}(\mathbf{x})|^2 d^3x$  diverges. Like position eigenstates, momentum eigenstates are also non-normalizable.

### 4.3 Rotations

We now consider the effect rotating  $\mathbb{R}^3$  has on a quantum state  $|\psi\rangle \in \mathcal{H}$ . For a vector  $\mathbf{v} \in \mathbb{R}^3$  a rotation anticlockwise around the axis  $\hat{\mathbf{a}}$  by an amount  $|\boldsymbol{\alpha}|$  is a linear transformation

$$\mathbf{R}(\boldsymbol{\alpha}) : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \quad \mathbf{R}(\boldsymbol{\alpha}) : \mathbf{v} \mapsto \mathbf{v}' = \mathbf{R}(\boldsymbol{\alpha})\mathbf{v} \quad (4.39)$$

that obeys

$$\mathbf{v}' \cdot \mathbf{v}' = \mathbf{v} \cdot \mathbf{v} \quad \text{and} \quad \det(\mathbf{R}(\boldsymbol{\alpha})) = +1. \quad (4.40)$$

The first of these conditions says that rotations preserve lengths, and implies that  $\mathbf{R}(\boldsymbol{\alpha})$  must be an *orthogonal* transformation. The second ensure that  $\mathbf{R}(\boldsymbol{\alpha})$  preserves the orientation. For infinitesimal rotations, figure 4 shows that

$$\mathbf{v}' = \mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v} + O(\delta\boldsymbol{\alpha}^2). \quad (4.41)$$

(Note that this obeys (4.40).) Given two orthogonal transformations  $\mathbf{R}(\boldsymbol{\alpha})$  and  $\mathbf{R}(\boldsymbol{\beta})$  the composite  $\mathbf{R}(\boldsymbol{\beta})\mathbf{R}(\boldsymbol{\alpha})$  is again an orthogonal transformation with unit determinant. Spatial rotations form a group, known as  $\text{SO}(3)$ , with the identity being the trivial rotation and the inverse of  $\mathbf{R}(\boldsymbol{\alpha})$  being a rotation of the same amount around the same axis  $\boldsymbol{\alpha}$ , but now clockwise. However, in general  $\mathbf{R}(\boldsymbol{\beta})\mathbf{R}(\boldsymbol{\alpha}) \neq \mathbf{R}(\boldsymbol{\alpha})\mathbf{R}(\boldsymbol{\beta})$ , so the order in which we apply rotations is important and the rotation group is *non-Abelian*.

As for any group of transformations, in quantum mechanics the group of rotations is represented on  $\mathcal{H}$  by unitary operators. We denote the unitary operator corresponding to  $\mathbf{R}(\boldsymbol{\alpha})$  as  $U(\boldsymbol{\alpha})$ .  $U(\boldsymbol{\alpha})$  acts on the position operator as

$$U^{-1}(\boldsymbol{\alpha}) \mathbf{X} U(\boldsymbol{\alpha}) = \mathbf{R}(\boldsymbol{\alpha}) \mathbf{X} \quad (4.42)$$

where the *lhs* involves composition of operators in  $\mathcal{H}$ , while the *rhs* here is the usual action of a rotation on the three components of  $\mathbf{X}$ . For example, if  $\boldsymbol{\alpha} = \alpha \hat{\mathbf{z}}$  so that the rotation is around the  $z$ -axis in  $\mathbb{R}^3$ , then in more detail (4.42) says

$$\begin{pmatrix} U^{-1}(\boldsymbol{\alpha}) X U(\boldsymbol{\alpha}) \\ U^{-1}(\boldsymbol{\alpha}) Y U(\boldsymbol{\alpha}) \\ U^{-1}(\boldsymbol{\alpha}) Z U(\boldsymbol{\alpha}) \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \quad (4.43)$$

Thus, conjugating the operator  $X$  by this rotation operator changes it to a new operator  $U^{-1}(\boldsymbol{\alpha}) X U(\boldsymbol{\alpha})$  on Hilbert space, which we identify with the usual linear combination  $X \cos \alpha - Y \sin \alpha$ . Again,  $U(\boldsymbol{\alpha})$  must act on  $\mathbf{X}$  this way if it is indeed to correspond to a rotation.

For an infinitesimal transformation, following (4.8) we can write

$$U(\delta \boldsymbol{\alpha}) = 1 - \frac{i}{\hbar} \delta \boldsymbol{\alpha} \cdot \mathbf{J} + O(\delta \boldsymbol{\alpha}^2) \quad (4.44)$$

for some Hermitian generators  $\mathbf{J}/\hbar$ . Since angles are dimensionless,  $\mathbf{J}$ , like  $\hbar$ , must have dimensions of (length  $\times$  momentum). Later we will see that  $\mathbf{J}$  corresponds to the angular momentum operator, though *by definition* we have that  $\mathbf{J}/\hbar$  is the generator of rotations. Using this and (4.41) in (4.42) shows that we have

$$\frac{i}{\hbar} [\delta \boldsymbol{\alpha} \cdot \mathbf{J}, \mathbf{X}] = \delta \boldsymbol{\alpha} \times \mathbf{X}, \quad (4.45)$$

and since this is true for any axis of rotation, we have the commutation relations

$$[J_i, X_j] = i\hbar \sum_k \epsilon_{ijk} X_k. \quad (4.46)$$

These relations are nothing more than the statement that the operator  $\mathbf{X}$  transforms as a vector under rotations. They are the infinitesimal version of (4.42) – they *must* hold if  $\mathbf{J}$  is indeed to generate rotations.

Just as for translations, the mutual commutation relations among the components of  $\mathbf{J}$  themselves follow from the fact that the  $U(\boldsymbol{\alpha})$ s provide a homomorphism from the group<sup>26</sup>

---

<sup>26</sup>Again, later we'll see that this statement needs to be slightly refined.

$SO(3)$  of rotations to the space of unitary operators on  $\mathcal{H}$ . However, the non-Abelian nature of the rotation group means we should expect these commutators to be non-trivial in general. Performing two infinitesimal rotations through angles  $\delta\boldsymbol{\alpha}$  and  $\delta\boldsymbol{\beta}$ , for any vector  $\mathbf{v}$  we have

$$\begin{aligned}\mathbf{R}(\delta\boldsymbol{\beta}) \mathbf{R}(\delta\boldsymbol{\alpha}) \mathbf{v} &= \mathbf{R}(\delta\boldsymbol{\beta})(\mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v}) = (\mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v}) + \delta\boldsymbol{\beta} \times (\mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v}) \\ &= \mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v} + \delta\boldsymbol{\beta} \times \mathbf{v} + \delta\boldsymbol{\beta} \times (\delta\boldsymbol{\alpha} \times \mathbf{v})\end{aligned}\quad (4.47)$$

to lowest non-trivial order in  $\delta\boldsymbol{\alpha}$  and  $\delta\boldsymbol{\beta}$ . Consequently, the difference between these rotations and the two rotations performed in the opposite order is

$$\begin{aligned}[\mathbf{R}(\delta\boldsymbol{\beta}) \mathbf{R}(\delta\boldsymbol{\alpha}) - \mathbf{R}(\delta\boldsymbol{\alpha}) \mathbf{R}(\delta\boldsymbol{\beta})] \mathbf{v} &= \delta\boldsymbol{\beta} \times (\delta\boldsymbol{\alpha} \times \mathbf{v}) - \delta\boldsymbol{\alpha} \times (\delta\boldsymbol{\beta} \times \mathbf{v}) \\ &= (\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) \times \mathbf{v} \\ &= \mathbf{R}(\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) \mathbf{v} - \mathbf{v}\end{aligned}\quad (4.48)$$

using standard properties of the vector triple product. The *rhs* again involves a rotation, through the angle  $\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}$ . Applying the homomorphism  $U$  we obtain

$$[U(\delta\boldsymbol{\beta}), U(\delta\boldsymbol{\alpha})] = U(\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) - I \quad (4.49)$$

for our operators in Hilbert space. Using (4.44), this is

$$-\frac{1}{\hbar^2} [\delta\boldsymbol{\beta} \cdot \mathbf{J}, \delta\boldsymbol{\alpha} \cdot \mathbf{J}] = -\frac{i}{\hbar} (\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) \cdot \mathbf{J} \quad (4.50)$$

and since this must hold for arbitrary successive infinitesimal rotations, we have finally

$$[J_i, J_j] = i\hbar \sum_k \epsilon_{ijk} J_k \quad (4.51)$$

as the commutation relations among the rotation generators.

Again, there's nothing ‘weird’ or ‘quantum’ about (4.51), beyond the fact that they involve operators on Hilbert space. In particular, their form just reflects the fact that the order of rotations around different axes matters, and that the difference between the two orderings is itself a rotation around the axis perpendicular to the original two. Compared to the relations  $[P_i, P_j] = 0$ , the non-triviality of the commutation relations (4.51) arises purely because  $SO(3)$  is a non-Abelian group, unlike the group of translations. These non-trivial commutation relations do not prevent us from exponentiating (4.44) to write  $U(\boldsymbol{\alpha}) = e^{-i\boldsymbol{\alpha} \cdot \mathbf{J}/\hbar}$  for a finite rotation around a fixed axis  $\boldsymbol{\alpha}$ , because the exponentiation always involved the same component  $\hat{\boldsymbol{\alpha}} \cdot \mathbf{J}$  of the rotation generator, which certainly commutes with itself.

The combined group of translations *and* rotations of Euclidean  $\mathbb{R}^3$  is sometimes known as  $E(3)$ , or else  $ISO(3)$ . For translations through  $a_1$  and  $a_2$ , and rotations  $R_1$  and  $R_2$ ,  $E(3)$  has the group composition law

$$(a_2, R_2) \cdot (a_1, R_1) = (a_2 + R_2 a_1, R_2 R_1), \quad (4.52)$$

where we note that the second rotation also acts on the first translation. Clearly, both  $\mathbb{R}^3$  and  $SO(3)$  are subgroups of  $E(3)$ , but the fact that rotations act non-trivially on previous translations means that  $E(3)$  is the *semi-direct* product  $E(3) \cong \mathbb{R}^3 \rtimes SO(3)$ . This group law shows that rotations and translations do not commute: translating any vector  $\mathbf{v}$  through  $\mathbf{a}$  and then rotating around  $\hat{\alpha}$  gives us  $\mathbf{R}(\alpha)(\mathbf{v} + \mathbf{a})$ , whereas first rotating then translating gives  $\mathbf{a} + \mathbf{R}(\alpha)\mathbf{v}$  instead. In particular, for infinitesimal translations and rotations we have

$$\mathbf{R}(\delta\alpha)(\mathbf{v} + \delta\mathbf{a}) - (\mathbf{R}(\delta\alpha)\mathbf{v} + \delta\mathbf{a}) = \delta\alpha \times \delta\mathbf{a} \quad (4.53)$$

independent of  $\mathbf{v}$ . Applying the homomorphism  $U$ , in Hilbert space we obtain commutation relations<sup>27</sup>

$$[U_R(\delta\alpha), U_T(\delta\mathbf{a})] = U_T(\delta\alpha \times \delta\mathbf{a}) - I \quad (4.54)$$

or equivalently

$$[J_i, P_j] = i\hbar \sum_k \epsilon_{ijk} P_k \quad (4.55)$$

in terms of the rotation and translation generators.

More generally, any 3-component linear operator  $\mathbf{V} : \mathcal{H} \rightarrow \mathcal{H}$  is said to *transform as a vector under rotations* if it obeys  $U^{-1}(\alpha)\mathbf{V}U(\alpha) = \mathbf{R}(\alpha)\mathbf{V}$  for any  $\alpha$ . (Depending on its behaviour under parity transformations, discussed in section 4.6, such a  $\mathbf{V}$  could either be a vector or pseudovector operator.) Just as above, for infinitesimal rotations this implies the commutation relations

$$[J_i, V_j] = i\hbar \sum_k \epsilon_{ijk} V_k \quad (4.56)$$

among the components of  $\mathbf{V}$  and  $\mathbf{J}$ . We see that  $\mathbf{X}$ ,  $\mathbf{P}$  and  $\mathbf{J}$  itself<sup>28</sup> each transform as vectors under rotations. This of course is just what we'd expect for position, momentum and angular momentum in classical mechanics.

On the other hand, if an operator  $S$  obeys

$$U^{-1}(\alpha) S U(\alpha) = S \quad (4.57)$$

for any  $\alpha$ , so that it is unchanged by the rotation operator, we say  $S$  is a *scalar* operator<sup>29</sup>. The corresponding infinitesimal version is

$$[\mathbf{J}, S] = 0. \quad (4.58)$$

Just as we can form a scalar by taking the Euclidean inner product  $\mathbf{v} \cdot \mathbf{w}$  of the two vectors  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ , so we can form scalar operators from the Euclidean inner product of two vector

---

<sup>27</sup>Here, for the sake of clarity, we have labelled the translations operator by  $U_T$  and rotation operator by  $U_R$ . In particular, the operator on the *rhs* of (4.54) is a translation.

<sup>28</sup>This is always true of  $\mathbf{X}$  and  $\mathbf{P}$ , but turns out to be a three-dimensional coincidence for  $\mathbf{J}$ : the rotation group  $SO(d)$  of  $\mathbb{R}^d$  has dimension  $d(d-1)/2$  and the generators are generically represented by antisymmetric matrices  $\mathcal{J}_{ij}$ . When  $d=3$ , a  $3 \times 3$  antisymmetric matrix has 3 independent components, so we can equivalently package the  $\mathcal{J}_{ij}$  as vectors through  $J_i = \epsilon_{ijk}\mathcal{J}_{jk}$ .

<sup>29</sup>More precisely, we say that  $S$  transforms as a scalar under rotations — depending on its behaviour under parity transformations it could be either a scalar or pseudoscalar operator.

operators. Indeed, if  $U^{-1}(\boldsymbol{\alpha})\mathbf{V}U(\boldsymbol{\alpha}) = \mathbf{R}(\boldsymbol{\alpha})\mathbf{V}$  and similarly for  $\mathbf{W}$ , then

$$\begin{aligned} U^{-1}(\boldsymbol{\alpha})(\mathbf{V} \cdot \mathbf{W})U(\boldsymbol{\alpha}) &= (U^{-1}(\boldsymbol{\alpha})\mathbf{V}U(\boldsymbol{\alpha})) \cdot (U^{-1}(\boldsymbol{\alpha})\mathbf{W}U(\boldsymbol{\alpha})) \\ &= (\mathbf{R}(\boldsymbol{\alpha})\mathbf{V}) \cdot (\mathbf{R}(\boldsymbol{\alpha})\mathbf{W}) \\ &= \mathbf{V} \cdot \mathbf{W} \end{aligned} \quad (4.59)$$

by the standard rotational invariance of the dot product. As always, we can write this in terms of commutators by considering infinitesimal rotations:

$$\begin{aligned} \left[ J_i, \sum_j V_j W_j \right] &= \sum_j [J_i, V_j] W_j + \sum_j V_j [J_i, W_j] \\ &= i\hbar \sum_{jk} \epsilon_{ijk} V_k W_j + i\hbar \sum_{jk} V_j \epsilon_{ijk} W_k \\ &= i\hbar \sum_{jk} \epsilon_{ijk} (-V_j W_k + V_j W_k) = 0, \end{aligned} \quad (4.60)$$

where in going to the final line we relabelled dummy indices  $j \leftrightarrow k$  in the first term and used the antisymmetry of  $\epsilon_{ijk}$ . Note that our calculations always preserved the order of  $\mathbf{V}$  and  $\mathbf{W}$  — our result holds irrespective of whether  $\mathbf{V}$  and  $\mathbf{W}$  commute.

An important special case of (4.58) is to take  $S = \mathbf{J}^2 = \mathbf{J} \cdot \mathbf{J}$ . The fact that the rotation generators obey the commutation relations

$$[J_i, J_j] = i\hbar \sum_k \epsilon_{ijk} J_k \quad \text{and} \quad [J_i, \mathbf{J}^2] = 0 \quad (4.61)$$

means that we cannot find a complete set of simultaneous eigenstates of all of the  $J_i$ s, but we *can* find a complete set of simultaneous eigenstates of  $\mathbf{J}^2$  and any one component of  $\mathbf{J}$ . Looking ahead to our interpretation of  $\mathbf{J}$  as the angular momentum operator, Born's 2<sup>nd</sup> postulate of QM tells us that we cannot say a particle has a definite angular momentum vector, but we can know the magnitude of the angular momentum and the amount aligned along any given axis.

#### 4.3.1 Translations Around a Circle

In 1B QM, you defined the orbital angular momentum operator  $\mathbf{L} = \mathbf{X} \times \mathbf{P}$  which also has the commutation relations (4.56) with itself and with  $\mathbf{X}$  and  $\mathbf{P}$ . We'll understand the relation between  $\mathbf{J}$  and  $\mathbf{L}$  later, but it's important to understand that in general  $\mathbf{J} \neq \mathbf{L}$ . We can understand  $\mathbf{L}$  from the present perspective as follows.

When a system is displaced through the vector  $\mathbf{a}$ , its state is transformed by the unitary operator  $U(\mathbf{a}) = e^{-i\mathbf{a} \cdot \mathbf{P}/\hbar}$ . We now imagine successively performing  $n$  translations successively through the set of vectors  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ . Each translation is represented by a unitary operator  $U(\mathbf{a}_i)$ , so the final state will be

$$U(\mathbf{a}_n) \cdots U(\mathbf{a}_2)U(\mathbf{a}_1)|\psi\rangle = U(\mathbf{b})|\psi\rangle \quad (4.62)$$

where  $\mathbf{b} = \mathbf{a}_1 + \dots + \mathbf{a}_n$  is the total displacement vector. Since the net result depends only on the total translation  $\mathbf{b}$ , the change in  $|\psi\rangle$  is independent of the particular path that the

system takes. In particular, if the path is closed so that  $\mathbf{a}_1 + \cdots + \mathbf{a}_n = \mathbf{0}$ , then  $|\psi\rangle$  is unchanged.

Now consider the effect of translating the system around an arc of a circle centred on the origin. We can approximate this circle by an  $N$ -sided regular polygon, with the approximation improving as  $N \rightarrow \infty$ , so we move around a circle by applying a succession of small translations, each in a slightly different direction. Specifically, if the system initially lies at some location  $\mathbf{x}$ , making an angle  $\alpha$  with some axis, then when move it in the plane normal to  $\mathbf{n}$  to lie at an angle  $\alpha + \delta\alpha$ , we translate it through  $\delta\mathbf{a} = \delta\alpha \mathbf{n} \times \mathbf{x}$ . Thus the associated unitary translation operator obeys

$$U^{-1}(\delta\mathbf{a}) \mathbf{X} U(\delta\mathbf{a}) = \mathbf{X} + \delta\alpha (\mathbf{n} \times \mathbf{X}) \quad (4.63)$$

and so can be written as

$$U(\delta\mathbf{a}) = 1 - \frac{i}{\hbar} \delta\alpha (\mathbf{n} \times \mathbf{X}) \cdot \mathbf{P} + O(\delta\alpha^2) = 1 - \frac{i}{\hbar} \delta\alpha \mathbf{n} \cdot \mathbf{L} + O(\delta\alpha^2), \quad (4.64)$$

where

$$\mathbf{L}/\hbar = \mathbf{X} \times \mathbf{P}/\hbar \quad (4.65)$$

is Hermitian operator. We now see that  $\mathbf{L}$  is the *generator of circular transformations*. (4.64) contains only one operator,  $\mathbf{n} \cdot \mathbf{L}$ , so it inevitably commutes with itself and we can exponentiate to find  $U(\mathbf{a}_{\text{circ}}) = e^{-i\alpha \mathbf{n} \cdot \mathbf{L}/\hbar}$  for finite translations around our circle.

#### 4.3.2 Spin

The commutation relations  $[L_i, X_j]$ ,  $[L_i, P_j]$  and  $[L_i, L_j]$  of the composite operator  $\mathbf{L}$  all follow from the more primitive commutation relations  $[X_i, P_j]$ ,  $[X_i, X_j]$  and  $[P_i, P_j]$ . If we're only concerned with these operators, then  $U_{\text{circ}} = e^{-i\alpha \mathbf{n} \cdot \mathbf{L}/\hbar}$  is indistinguishable from the rotation operator  $U(\boldsymbol{\alpha}) = e^{-i\boldsymbol{\alpha} \cdot \mathbf{J}/\hbar}$ . However, if our system has any internal structure, circular translations using the centre of mass position operator  $\mathbf{X}$  and centre of mass



**Figure 5:** The rotation operator  $U(\boldsymbol{\alpha})$  swings a system around the origin and also rotates its orientation in  $\mathbb{R}^3$ , while a circular translation merely moves the system around a circular path, without affecting its orientation. The difference is a rotation of the body around its own centre of mass, reorientating it without changing its location.

momentum  $\mathbf{P}$  are not the same as rotations: the difference is best explained by a picture, which you can find in figure 5.

We define the *spin operator*  $\mathbf{S}$  to be the difference

$$\mathbf{S} = \mathbf{J} - \mathbf{L} \quad (4.66)$$

so that  $\mathbf{J} = \mathbf{L} + \mathbf{S}$ . From figure 5 we see that difference between rotating an object around some fixed origin and a translating it's centre of mass along a circular path is a rotation around the object's centre of mass. generates a rotation of the body *around its own centre of mass*. We thus expect (and will confirm below) that  $\mathbf{S}$  generates rotations of a body around its own centre of mass, reorienting it in space. This is why  $\mathbf{S}$  is called the spin operator.

In the case of a macroscopic body, made up of many constituent particles, it's reasonable to suppose that the spin operator just account for the difference between translating the body as a whole and translating each individual particle around their own arcs, with slightly different radii according to where in the body the particle is located. That is,

$$\mathbf{S} \stackrel{?}{=} \left( \sum_a \mathbf{X}_a \times \mathbf{P}_a \right) - \mathbf{X} \times \mathbf{P} = \sum_a \mathbf{x}_a \times \mathbf{p}_a \quad (4.67)$$

where  $\mathbf{X}_a$  and  $\mathbf{P}_a$  are position and translation operators for each individual particle and  $\mathbf{x}_a$  and  $\mathbf{p}_a$  are their positions and momenta *relative* to the centre of mass<sup>30</sup>. However, for an object consisting of many particles such a description is clearly going to be very cumbersome. More fundamentally, *we do not know what the ‘fundamental’ constituents of our object really are*. (For example, if I sit on a merry-go-round, is the ‘right’ quantum description of my motion given in terms of my cells, or my atoms, or protons and neutrons, or quarks and gluons, or bits of string, or ...?) It's thus crucial that we can consider objects *as a whole* in quantum mechanics, just as we can classically. For rotations, understanding  $\mathbf{S}$  will allow us to do this. In addition, as we'll see in section 5.3.1, one of the surprises of quantum mechanics is that even fundamental particles such as electrons and photons may have an ‘intrinsic’ spin that (as far as we know) is not related to any composite structure.

Fortunately, commutation relations involving the spin operator  $\mathbf{S}$  are easy to obtain from the ones we already have for  $\mathbf{J}$  and  $\mathbf{L}$ . Firstly, the fundamental relations

$$[J_i, X_j] = i\hbar \sum_k \epsilon_{ijk} X_k \quad \text{and} \quad [J_i, P_j] = i\hbar \sum_k \epsilon_{ijk} P_k \quad (4.68)$$

show that

$$[J_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k, \quad (4.69)$$

---

<sup>30</sup>We'll understand how to properly describe the quantum mechanics of a system with many degrees of freedom (*e.g.* many particles) in chapter 7.

so that the angular momentum operator  $\mathbf{L}$  also transforms as a vector under rotations. Recalling from 1B QM that<sup>31</sup>  $[L_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k$ , we have

$$\begin{aligned}[S_i, S_j] &= [J_i - L_i, J_j - L_j] = [J_i, J_j] - [J_i, L_j] - [L_i, J_j] + [L_i, L_j] \\ &= i\hbar \sum_k \epsilon_{ijk} (J_k - L_k) = i\hbar \sum_k \epsilon_{ijk} S_k.\end{aligned}\tag{4.70}$$

It then immediately follows that  $[\mathbf{S}, \mathbf{S}^2] = 0$ . This algebra – the same as that obeyed by the components of  $\mathbf{J}$  and  $\mathbf{L}$  – confirms that the spin operators  $\mathbf{S}$  generate some form of rotation. On the other hand, since  $[L_i, X_j] = i\hbar \sum_k \epsilon_{ijk} X_k$  and  $[L_i, P_j] = i\hbar \sum_k \epsilon_{ijk} P_k$ , we have

$$\begin{aligned}[S_i, X_j] &= [J_i, X_j] - [L_i, X_j] = 0 \\ [S_i, P_j] &= [J_i, P_j] - [L_i, P_j] = 0\end{aligned}\tag{4.71}$$

These commutation relations confirm that the spin operator has nothing to do with an object's location in or motion through space, but is purely to do with rotating its *intrinsic* orientation. Finally, since  $\mathbf{S}$  commutes with both  $\mathbf{X}$  and  $\mathbf{P}$ , it also commutes with  $\mathbf{L}$ :

$$[S_i, L_j] = 0 \quad \forall i, j.\tag{4.72}$$

This allows us to factorize the operator  $U(\boldsymbol{\alpha})$  describing finite rotations as

$$U(\boldsymbol{\alpha}) = e^{-i\boldsymbol{\alpha} \cdot \mathbf{J}/\hbar} = e^{-i\boldsymbol{\alpha} \cdot (\mathbf{L} + \mathbf{S})/\hbar} = e^{-i\boldsymbol{\alpha} \cdot \mathbf{L}/\hbar} e^{-i\boldsymbol{\alpha} \cdot \mathbf{S}/\hbar} = e^{-i\boldsymbol{\alpha} \cdot \mathbf{S}/\hbar} e^{-i\boldsymbol{\alpha} \cdot \mathbf{L}/\hbar}.\tag{4.73}$$

As in figure 5, these equations confirm that we can think of a quantum rotation as consisting of a translation of a body's centre of mass along an arc centred on the origin together with a simultaneous rotation of the body around its own centre of mass by the same amount. The order in which we perform these two operations makes no difference.

#### 4.4 Time Translations

It's not only transformations of space that are represented on  $\mathcal{H}$  by unitary operators. Consider translations in time. These again form an Abelian group, since sending  $t_0$  to  $t_0 + t$  and then to  $t_0 + t + t'$  gives the same end result as does  $t_0 \rightarrow t_0 + t' \rightarrow t_0 + t' + t$ . If we want the total probability of our particle being found *somewhere* in  $\mathbb{R}^3$  at *all* times, then time translation must also be represented by a unitary operator  $U(t) : \mathcal{H} \rightarrow \mathcal{H}$ . Since we can translate through an arbitrarily small time, the time translation operator takes the form

$$U(t) = \exp\left(-\frac{i}{\hbar} H t\right)\tag{4.74}$$

where  $H/\hbar$  must be Hermitian and is, by definition, the generator of time translations. This generator must have dimensions  $1/(\text{time})$ , so the conventional factor of  $\hbar$  means that  $H$  itself has dimensions of energy. Of course,  $H$  will turn out to be the Hamiltonian, but for now I'd like you to think of it more abstractly just as the generator of time translations.

---

<sup>31</sup>Check you're able to derive this statement from the definition of  $\mathbf{L}$  and the fundamental commutation relations  $[X_i, P_j] = i\hbar \delta_{ij}$ .

The fact that  $U(t)$  represents the effect of a time translation on  $\mathcal{H}$  means that if a particle is in some state  $|\psi(0)\rangle$  at time  $t_0 = 0$ , translating forward to time  $t$  we will find it in the state

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle. \quad (4.75)$$

In particular, the difference between  $|\psi(t)\rangle$  and the state we find a short time later is

$$|\psi(t + \delta t)\rangle - |\psi(t)\rangle = -\frac{i}{\hbar} \delta t H |\psi(t)\rangle + O(\delta t^2) \quad (4.76)$$

or, taking the limit  $\delta t \rightarrow 0$ ,

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle. \quad (4.77)$$

This, famously, is the *Time Dependent Schrödinger Equation*, which I'll often abbreviate to *TDSE* in what follows. It's just the infinitesimal version of the statement that all states in  $\mathcal{H}$  evolve in time according to the action of a unitary operator  $U(t)$ .

#### 4.4.1 The Heisenberg Picture

In 1B QM, we were used to the idea that states evolve in time according to the TDSE. However, just as we did for spatial transformations, we can instead work in a picture where the states are time independent and instead the *operators* evolve in time. To see how this works, suppose  $O_S$  is some operator that contains no explicit time dependence. Then the amplitude for the state  $O_S|\psi(t)\rangle$  to agree at time  $t$  with the state  $|\chi(t)\rangle$  is

$$\langle \chi(t) | O_S | \psi(t) \rangle = \langle \chi(0) | U^{-1}(t) O_S U(t) | \psi(0) \rangle \quad (4.78)$$

The right hand side here only makes explicit reference to the initial values of the states  $|\psi\rangle$  and  $|\chi\rangle$ . Since it holds true for any pair of states, just as above we can obtain the same results by always working with the initial states and instead evolving the operators as

$$O_H(t) = U^{-1}(t) O_S U(t). \quad (4.79)$$

The version of quantum mechanics where we use the TDSE to evolve the states in time, leaving operators unaltered, is known as the *Schrödinger picture*, whereas the version where the states are fixed at their initial values and instead the operators evolve in time is called the *Heisenberg picture*.

In fact, all this has a precise analogue in classical mechanics. Classically, there are also two ways of thinking about time evolution. On the one hand, we can think of a particle moving in some way through phase space  $M$ . If we know its location  $(\mathbf{x}(t), \mathbf{p}(t)) \in M$  for every time  $t$  we can compute any quantity we wish, represented by some function  $f : M \rightarrow \mathbb{R}$ , by evaluating  $f$  at the location of our particle, obtaining the value  $f(\mathbf{x}(t), \mathbf{p}(t))$ . (This is the perspective we took in the Introduction.) However, Newton's Laws are deterministic, so, given a force, the entire trajectory is determined by the initial conditions  $(\mathbf{x}_0, \mathbf{p}_0)$ . This suggests a perspective in which the ‘state’ of our particle is simply a choice of initial conditions. These initial conditions do not themselves evolve, rather, it is the quantities we measure that vary in time. Thus, instead of thinking of a physical quantity  $f$  as a map from

phase space, we treat it just as a map from time, so  $f : [t_0, \infty) \rightarrow \mathbb{R}$ . You'll examine these classical pictures further if you're taking the Classical Dynamics course, in the context of Hamiltonian mechanics. (It's really for this reformulation of classical mechanics, not the particular  $H$ , that Hamilton is famous.)

Differentiating (4.79) wrt  $t$  shows that

$$\begin{aligned} \frac{d}{dt} O_H(t) &= \frac{d}{dt} (U^{-1}(t) O_S U(t)) = \frac{i}{\hbar} (U^{-1}(t) H O_S U(t) - U^{-1} O_S H U(t)) \\ &= \frac{i}{\hbar} [H, O_H(t)], \end{aligned} \quad (4.80)$$

where in the last step we used the fact that  $[U(t), H] = 0$ , since  $U(t)$  depends only on  $H$ . This is *Heisenberg's equation of motion*. It's completely equivalent to Schrödinger's equation (4.77), and is also just a (very important!) special case of our general argument of how operators transform under the action of unitary groups of transformations. More generally, if the original operator  $O_S$  had some explicit time dependence of its own — independent of any particular particle's motion — then we would obtain a further term in this equation, modifying it to

$$\frac{d}{dt} O_H(t) = U^{-1}(t) \frac{\partial O_S}{\partial t} U(t) + \frac{i}{\hbar} [H, O_H(t)]. \quad (4.81)$$

For example, we may wish to understand the behaviour of a charged particle in the presence of an applied electric field. If the electric field is itself changing in time, then the potential in which the particle moves will be time-dependent even in the Schrödinger picture.

## 4.5 Dynamics

So far, we've simply defined some unitary operators  $U(\mathbf{a})$ ,  $U(\boldsymbol{\alpha})$  and  $U(t)$  that are responsible for translating, rotating or evolving our system through space and time. The commutation relations of the generators  $\mathbf{P}$  and  $\mathbf{J}$ , and their commutation relations with  $\mathbf{X}$  were determined purely by the properties of the corresponding group of transformations of  $\mathbb{R}^3$ . However, while we've seen that commutators such as  $[H, \mathbf{X}]$ ,  $[H, \mathbf{P}]$  and  $[H, \mathbf{J}]$  are important in the Heisenberg picture, telling us how these operators change in time, we haven't yet given any way to actually *calculate* what such commutators should be. Similarly, although we've said that in the Schrödinger picture, states evolve in time according to  $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ , we haven't given any way to work out what the action of the generator  $H$  on a state should actually be.

To do so, we must specify the *dynamics*: we must provide a relation  $H = H(\mathbf{X}, \mathbf{P})$  giving  $H$  in terms of the other operators whose commutators we already understand. The simplest (non-trivial) such relation is<sup>32</sup>

$$H = \frac{1}{2m} \mathbf{P}^2 \quad (4.82)$$

---

<sup>32</sup>Note that this is the *first* time we've pinned ourselves down to a non-relativistic theory. Up to this point, everything we've said holds good in relativistic quantum mechanics and, suitably interpreted, also in quantum field theory. In the lectures we studied translations and rotations, but in the **first problem set** you'll extend this to also consider Galilean boosts and will show that the non-relativistic Hamiltonian is compatible with the Galilean algebra. If one instead wishes to study a relativistic version of quantum theory, one starts by finding unitary operators representing the action of the Poincaré group  $\mathbb{R}^{3,1} \rtimes SO(3, 1)$

where  $m$  is a constant with dimensions of mass. Equation (4.82) relates how a state evolves in time (via  $H$ ) to how its location is translated through space (via  $\mathbf{P}$ ); in other words, it's telling us something about how particles *move!* In particular, with this form of Hamiltonian the difference between the expected location of a state  $|\psi\rangle$  at an initial time and a short time  $\delta t$  later is

$$\begin{aligned}\langle\psi|U^{-1}(\delta t)\mathbf{X}U(\delta t)|\psi\rangle - \langle\psi|\mathbf{X}|\psi\rangle &= \frac{i}{\hbar}\delta t\langle\psi|[H,\mathbf{X}]|\psi\rangle + O(\delta t^2) \\ &= \frac{\delta t}{m}\langle\psi|\mathbf{P}|\psi\rangle + O(\delta t^2).\end{aligned}\quad (4.83)$$

In the limit  $\delta t \rightarrow 0$  we see that the expected *velocity* of the particle is  $\langle\psi|\mathbf{P}|\psi\rangle/m$ . Equivalently, since this is true for all  $|\psi\rangle \in \mathcal{H}$ , we can write

$$\frac{d}{dt}\mathbf{X}(t) = \frac{\mathbf{P}(t)}{m}\quad (4.84)$$

in the Heisenberg picture. Furthermore, since  $[H, \mathbf{P}] = 0$  for this Hamiltonian,  $\mathbf{P}(t) = \mathbf{P}(0)$  and all states travel in uniform motion.

In the real world, we observe that particles do not always travel with constant velocities: they may slow down, speed up, or change direction as they encounter various obstacles. These obstacles are typically located at various different points in space. To allow for this, we generalise our dynamical relation (4.82) to

$$H = \frac{1}{2m}\mathbf{P}^2 + V(\mathbf{X}).\quad (4.85)$$

The first term on the right is the *kinetic term* and is the contribution to the Hamiltonian due to the state's travelling through space. The second, *potential term* is the contribution due to the state's location. This familiar form is still a very special case of the general statement  $H = H(\mathbf{X}, \mathbf{P})$  and later in the course we'll meet examples of Hamiltonians that don't fit this form.

Repeating the calculation of (4.84) we still find that  $d\mathbf{X}(t)/dt = \mathbf{P}(t)/m$ , but now  $[H, \mathbf{P}] \neq 0$  so the rate of motion through space will not always be the same. Rather, using (4.80), the Heisenberg picture operators obey

$$\frac{d}{dt}\mathbf{P}(t) = \frac{i}{\hbar}[H, \mathbf{P}(t)] = -\nabla V(t),\quad (4.86)$$

where  $\nabla V(t) = U^{-1}(t)(\partial V/\partial\mathbf{X})U(t) = -\nabla V(\mathbf{X}(t))$ . Thus the motion slows down or speeds up according to the gradient of  $V(\mathbf{X})$ .

Most of our common intuition about momentum comes from equation (4.86). We 'feel' that a tennis ball travelling with a certain speed has less momentum than a cannonball travelling with the same speed, and less than the same tennis ball travelling faster, because

---

of special relativity. The relation (4.82) between  $H$  and  $\mathbf{P}$  does not respect the commutation relations appropriate for Poincaré symmetry, but of course the dynamical relation  $H^2 = c^2\mathbf{P}^2 + m^2c^4$  does. The difficulty with relativistic quantum mechanics lies not with symmetries, but with interactions. The proper treatment of these requires Quantum Field Theory, a far deeper subject.

we've known from an early age that it will cause us less damage to stand in the way of the first tennis ball than either of the other two. This is really a statement about what our unfortunate bodies will have to do in order to bring these projectiles to rest (or how much effort we will have to exert to launch them in the first place). In other words, our intuitive notion of momentum is built on a feeling for the energy our body will gain as we slow the projectile down during the impact. This energy then excites the atoms that ultimately make up our nerves, muscles and bones, becoming dissipated through our bodies. Ultimately, it is the dynamical relation between the Hamiltonian and other operators that justifies us identifying the operator  $\mathbf{P}/\hbar$  — by definition, the generator of translations — with our pre-existing notion of *momentum* (in units of  $\hbar$ ). Only after we specify our dynamics do commutation relations such as  $[X_i, P_j] = i\hbar \delta_{ij}$  tell us, via Born's 2<sup>nd</sup> postulate of QM, that a particle cannot simultaneously be in a state of well-defined position and of well-defined momentum.

We'll often find it useful to write the kinetic term in a slightly different form. Keeping careful track of the operator ordering, you'll show in the problem sets that

$$\mathbf{L}^2 = (\mathbf{X} \times \mathbf{P}) \cdot (\mathbf{X} \times \mathbf{P}) = \mathbf{X}^2 (\mathbf{P}^2 - P_r^2) , \quad (4.87)$$

where for  $\hat{\mathbf{X}} = \mathbf{X}/|\mathbf{X}|$ ,

$$P_r = \frac{1}{2} (\hat{\mathbf{X}} \cdot \mathbf{P} + \mathbf{P} \cdot \hat{\mathbf{X}}) \quad (4.88)$$

is the radial momentum operator. Consequently, we can write the Hamiltonian (4.85) in terms of the generator of circular translations and the radial momentum operator  $P_r$  as

$$H = \frac{P_r^2}{2m} + \frac{1}{2m\mathbf{X}^2} \mathbf{L}^2 + V(\mathbf{X}) \quad (4.89)$$

where we note that  $[\mathbf{L}, \mathbf{X}^2] = 0$ , so the order of the operators in the second term is irrelevant provided we keep  $\mathbf{X}^2$  together as a composite operator. (In particular, in the position representation,  $\mathbf{X}^2$  acts as multiplication by  $\mathbf{x}^2 = r^2$ , which is certainly unaffected by translation in a circle.) Most of our intuitive feeling about angular momentum — mine comes from happy childhood hours spent on merry-go-rounds — is encapsulated in this relation between the Hamiltonian and  $\mathbf{L}^2$ . Again, it's the form of the Hamiltonian which is the basis of our identification of the generator  $\mathbf{L}/\hbar$  of circular transformations with angular momentum (in units of  $\hbar$ ).

#### 4.5.1 Symmetries and Conservation Laws

One of the immediate uses of the Heisenberg picture is to deduce a relation between symmetries of the Hamiltonian and conserved quantities<sup>33</sup>. If it happens that an operator  $Q$  commutes with the Hamiltonian, then it also commutes with  $e^{-iHt/\hbar}$ , so the Heisenberg picture operator

$$Q(t) = U^{-1}(t) Q U(t) = U^{-1}(t) U(t) Q = Q , \quad (4.90)$$

---

<sup>33</sup>Those of you taking the Classical Dynamics course will see a corresponding relation in Hamilton's approach to classical physics.

coinciding with the Schrödinger operator for all  $t$ . Infinitesimally, provided  $Q$  has no explicit time dependence, this is

$$\frac{d}{dt}Q(t) = \frac{i}{\hbar}[H, Q(t)] = \frac{i}{\hbar}U^{-1}(t)[H, Q]U(t) = 0. \quad (4.91)$$

Operators that are time independent even in the Heisenberg picture are said to be *conserved*. We've shown that conserved operators are just those that commute with the Hamiltonian.

Suppose we prepare a particle to be in an eigenstate of some conserved operator at time  $t = 0$ , so  $Q|\psi(0)\rangle = q|\psi(0)\rangle$ . Then at any later time we have

$$Q|\psi(t)\rangle = QU(t)|\psi(0)\rangle = U(t)Q|\psi(0)\rangle = qU(t)|\psi(0)\rangle = q|\psi(t)\rangle \quad (4.92)$$

so our particle remains in an eigenstate of  $Q$  at all subsequent times (provided the state evolves according to the TDSE). For this reason, it's usually sensible to expand our states in a basis of eigenstates of a maximal set of *conserved* operators, rather than a maximal commuting set of any old operators. We label the states in this basis by their corresponding eigenvalues, because the same labelling will remain valid at subsequent times. For example, provided  $H$  has no explicit time dependence<sup>34</sup>  $[H, H] = 0$  trivially, so the Hamiltonian itself is always conserved, and it is often useful to work in a basis of energy eigenstates.

By far the most important source of conserved quantities is *symmetries* of the Hamiltonian. These are transformations that leave the Hamiltonian invariant. We've seen that a generic unitary operator  $U(\theta) = e^{-i\theta T}$  representing some transformation of space on  $\mathcal{H}$  acts on operators as in equation (4.14). In particular, its action on the Hamiltonian is

$$H \mapsto U^{-1}(\theta)HU(\theta). \quad (4.93)$$

If the Hamiltonian is invariant under this transformation – *i.e.*, if the transformation is a symmetry – then

$$U^{-1}(\theta)HU(\theta) = H, \quad (4.94)$$

or equivalently

$$[T, H] = 0 \quad (4.95)$$

for an infinitesimal symmetry transformation, where  $T$  is the Hermitian generator. This is exactly the same condition (4.91) we had for the generator  $T$  to be conserved, so *symmetries of the Hamiltonian correspond to conserved quantities*<sup>35</sup>.

For example, if the Hamiltonian is translationally invariant then  $[\mathbf{P}, H] = 0$ , so for such Hamiltonians, momentum will be conserved. Note that it's not necessary for the Hamiltonian to be independent of each particle's position  $\mathbf{x}_a$  if  $H$  is to be translationally

<sup>34</sup>Even if the Hamiltonian does contain explicit time dependence, for example because it describes the dynamics of a charged particle in a varying electric field,  $[H(t), H(t)] = 0$  trivially. However,  $[H(t'), H(t)]$  may not vanish as the form of the Hamiltonian itself changes.

<sup>35</sup>If you're taking the Classical Dynamics or Integrable Systems courses, you'll meet the corresponding statement in classical mechanics in the guise of Nöther's theorem.

invariant, so long as any potential term  $V(\mathbf{x}_a - \mathbf{x}_b)$  depends only on the relative positions. Similarly, if the Hamiltonian is rotationally invariant then  $[\mathbf{J}, H] = 0$  so angular momentum will be conserved. In this case,  $J_z$ ,  $\mathbf{J}^2$  and  $H$  form a set of mutually commuting operators, so we can expand a general state in a basis  $\{|n, \ell, m, \dots\rangle\}$  where the labels  $(n, \ell, m)$  refer to the eigenvalues of  $H$ ,  $\mathbf{J}^2$  and  $J_z$ . We've left open the possibility that our states have further conserved properties, indexed by further labels. These will often contain information about the ‘internal’ state of our system.

## 4.6 Parity

Not all transformations are continuous. In physics, the most prominent example of a discrete transformation is the *parity transformation*  $\mathcal{P}$ , acting on  $\mathbb{R}^3$  as  $\mathcal{P} : \mathbf{x} \mapsto -\mathbf{x}$ . Since  $\det(\mathcal{P}) = -1$ , this is different from a rotation, so may have consequences that cannot be deduced by considering only rotations.

On Hilbert space, parity transformations will still be represented by a unitary operator, but this operator will no longer involve a parameter that may be taken infinitesimally small. Consequently there's no generator associated with parity transformations. Calling the unitary parity operator  $\Pi$ , we have

$$\Pi^{-1} \mathbf{X} \Pi = \mathcal{P} \mathbf{X} = -\mathbf{X} \quad (4.96)$$

as its defining property. Multiplying through on the right by  $\Pi$ , we can also write this as

$$\{\Pi, \mathbf{X}\} = \Pi \mathbf{X} + \mathbf{X} \Pi = 0. \quad (4.97)$$

where the bracket  $\{A, B\}$  is sometimes called the *anticommutator* of  $A$  and  $B$ . Translating through  $\mathbf{a}$  and then applying the parity operator  $\mathcal{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  is the same as first applying  $\mathcal{P}$ , then translating through  $-\mathbf{a}$ , so

$$\Pi^{-1} U(\mathbf{a}) \Pi = U(-\mathbf{a}), \quad (4.98)$$

so that the momentum operator also anticommutes with the parity operator:  $\Pi^{-1} \mathbf{P} \Pi = -\mathbf{P}$ . More generally, if  $\mathbf{V}$  is any operator that transforms as a vector under rotations, we say  $\mathbf{V}$  is a *vector operator* if also

$$\Pi^{-1} \mathbf{V} \Pi = -\mathbf{V}, \quad (4.99)$$

so that  $\{\Pi, V\} = 0$ .

If instead

$$\Pi^{-1} \mathbf{V} \Pi = +\mathbf{V} \quad (4.100)$$

or equivalently  $[\Pi, \mathbf{V}] = 0$ , then  $\mathbf{V}$  is a *pseudovector operator* (provided it transforms appropriately under rotations). The most prominent example of a pseudovector operator is the rotation generator  $\mathbf{J}$ : since the parity transformation  $\mathcal{P}$  acts on  $\mathbb{R}^3$  as  $-I_{3 \times 3}$ , rotations obey  $\mathbf{R}(\boldsymbol{\alpha}) \mathcal{P} = \mathcal{P} \mathbf{R}(\boldsymbol{\alpha})$ , or  $\mathcal{P}^{-1} \mathbf{R}(\boldsymbol{\alpha}) \mathcal{P} = \mathbf{R}(\boldsymbol{\alpha})$ . This gives

$$\Pi^{-1} U(\boldsymbol{\alpha}) \Pi = U(\boldsymbol{\alpha}) \quad (4.101)$$

for the parity and rotation operators in Hilbert space, or  $\Pi^{-1}\mathbf{J}\Pi = \mathbf{J}$  for the rotation generator. Likewise, the orbital angular momentum operator obeys

$$\Pi^{-1}\mathbf{L}\Pi = \Pi^{-1}(\mathbf{X} \times \mathbf{P})\Pi = (\Pi^{-1}\mathbf{X}\Pi) \times (\Pi^{-1}\mathbf{P}\Pi) = (-)^2 \mathbf{X} \times \mathbf{P} = +\mathbf{L} \quad (4.102)$$

so is a pseudovector as in classical mechanics. It follows that the spin operator  $\mathbf{S} = \mathbf{J} - \mathbf{L}$  is also a pseudovector.

Similarly, operators  $S$  that are invariant under rotations are *scalar operators* if in addition  $\Pi^{-1}S\Pi = S$  so that they are unchanged under parity, and are *pseudoscalar operators* if instead  $\Pi^{-1}S\Pi = -S$ . (Notice that the signs for scalars and pseudoscalars are the opposite way round to those for vectors and pseudovectors.) As usual, the parity of a system will be conserved if  $[H, \Pi] = 0$ . As a simple example, if our system is governed by the dynamical relation  $H = \mathbf{P}^2/2m + V(\mathbf{X})$  where the potential is an even function, then

$$\begin{aligned} \Pi^{-1}H\Pi &= \frac{1}{2m}(\Pi^{-1}\mathbf{P}\Pi) \cdot (\Pi^{-1}\mathbf{P}\Pi) + \Pi^{-1}V(\mathbf{X})\Pi \\ &= \frac{\mathbf{P}^2}{2m} + V(\Pi^{-1}\mathbf{X}\Pi) \\ &= \frac{\mathbf{P}^2}{2m} + V(\mathbf{X}) = H, \end{aligned} \quad (4.103)$$

so parity is conserved for such Hamiltonians.

Parity transformations form the group  $G = \mathbb{Z}_2$ , since carrying out a non-trivial parity transformation twice just brings us back to where we were. Thus  $\mathcal{P}^2 = 1_{\mathbb{R}^3}$  and applying our homomorphism shows that

$$\Pi^2 = 1_{\mathcal{H}} \quad (4.104)$$

for the parity operator on Hilbert space, too. Now suppose that  $|\psi\rangle$  is an eigenstate of  $\Pi$  with eigenvalue  $\eta$ . Then<sup>36</sup>

$$|\psi\rangle = \Pi^2|\psi\rangle = \eta\Pi|\psi\rangle = \eta^2|\psi\rangle \quad (4.105)$$

so the spectrum of  $\Pi$  is just  $\{+1, -1\}$ .

If  $|\mathbf{v}\rangle$  is an eigenstate of a vector operator  $\mathbf{V}$ , with  $\mathbf{V}|\mathbf{v}\rangle = \mathbf{v}|\mathbf{v}\rangle$ , then from (4.99), the parity reversed state  $|\mathbf{v}'\rangle = \Pi|\mathbf{v}\rangle$  obeys

$$\mathbf{V}|\mathbf{v}'\rangle = \mathbf{V}\Pi|\mathbf{v}\rangle = -\Pi\mathbf{V}|\mathbf{v}\rangle = -\mathbf{v}\Pi|\mathbf{v}\rangle = -\mathbf{v}|\mathbf{v}'\rangle \quad (4.106)$$

so is also an eigenstate of  $\mathbf{V}$ , but with opposite sign eigenvalue. In particular,  $\Pi|\mathbf{x}\rangle$  is an eigenstate of the position operator with eigenvalue  $-\mathbf{x}$ , so

$$\Pi|\mathbf{x}\rangle = c|-\mathbf{x}\rangle. \quad (4.107)$$

for some constant  $c \in \mathbb{C}$ . This is not an eigenvalue equation, because the *rhs* involves  $|-\mathbf{x}\rangle$  rather than  $|\mathbf{x}\rangle$ . However, we can immediately identify  $c$  with  $\eta$  because applying  $\Pi$  a second time gives<sup>37</sup>

$$|\mathbf{x}\rangle = \Pi^2|\mathbf{x}\rangle = c^2|\mathbf{x}\rangle, \quad (4.108)$$

---

<sup>36</sup>This argument is adequate for our course, and leads to a correct conclusion, but it hides a number of subtleties. See *e.g.* Weinberg *The Quantum Theory of Fields*, vol. 1, section 4.7 for a fuller discussion.

<sup>37</sup>Note that it is important here that  $\Pi|\mathbf{x}\rangle = c|-\mathbf{x}\rangle$  for *all* values of  $\mathbf{x}$ .

so again  $c^2 = 1$  and  $c = \pm 1$ . More generally, given any state  $|\psi\rangle$ , the wavefunction of the parity transformed state  $\Pi|\psi\rangle$  is

$$\langle \mathbf{x} | \Pi | \psi \rangle = \eta^{-1} \langle -\mathbf{x} | \psi \rangle = \eta \psi(-\mathbf{x}), \quad (4.109)$$

so, up to a sign, the wavefunction of the new state takes the same value at  $\mathbf{x}$  as the original one did at  $-\mathbf{x}$ .

For example, suppose that for some particle,  $|\mathbf{x}\rangle$  obeys  $\Pi|\mathbf{x}\rangle = |-\mathbf{x}\rangle$ , with  $\eta = +1$ . Recalling from 1B QM that the spherical harmonics  $Y_\ell^m(\mathbf{x})$  obey  $Y_\ell^m(-\mathbf{x}) = (-1)^\ell Y_\ell^m(\mathbf{x})$ , we see that if this particle is in the state  $|n, \ell, m\rangle$  whose wavefunction  $\langle \mathbf{x} | n, \ell, m \rangle = R_n(|\mathbf{x}|) Y_\ell^m(\mathbf{x})$  (where the radial part  $R_n(r)$  of the wavefunction determines the energy level  $n$ ), applying the parity operator gives

$$\langle \mathbf{x} | \Pi | n, \ell, m \rangle = \langle -\mathbf{x} | n, \ell, m \rangle = (-1)^\ell \langle \mathbf{x} | n, \ell, m \rangle \quad (4.110)$$

and consequently  $\Pi|n, \ell, m\rangle = (-1)^\ell |n, \ell, m\rangle$ .

Parity transformations act in a more complicated way on macroscopic objects, or systems with some internal structure. This is clear classically: the parity transformation of a book initially located at  $\mathbf{x}$  is not simply a book located at  $-\mathbf{x}$ , but rather a mirror image (simultaneously left-right, up-down and front-back) of the book. As with the distinction between  $\mathbf{J}$  and  $\mathbf{L}$ , for a macroscopic body we might hope to account for this by defining a separate parity operator for each constituent particle. However, it turns out that subatomic particles such as the photon or pion have their own ‘intrinsic’ parities. Thus, whether a given  $\Pi$ -eigenstate has  $\eta = +1$  or  $\eta = -1$  depends not only on details of the spatial wavefunction, but also on the species of particle the state describes.

We’ll explore intrinsic parity further in section 7.4, but as an example of its use, consider transitions between different energy levels of an atom. These are usually mediated by electromagnetic interactions: during such a transition, the atom’s electrons either emit or absorb a photon whose energy  $\hbar\omega$  is equal to the difference in energy  $\Delta E$  between the two levels involved in the transition. As we’ll study in more detail later, when the corresponding wavelength is much larger than the typical size of the atom, the transition rate  $\Gamma(i \rightarrow f)$  is given by

$$\Gamma(i \rightarrow f) = \frac{4(\Delta E)^3}{c^3 \hbar^4} |\langle f | \mathbf{D} | i \rangle|^2, \quad (4.111)$$

where  $|i\rangle$  and  $|f\rangle$  are the initial and final states of the atom, and  $\mathbf{D} = \sum e_a \mathbf{X}_a$  is the operator corresponding to the atom’s electric dipole moment. (The sum runs over all the electrons in the atom.) Reversing the parity of space affects all the electrons equally, so the parity operator obeys  $\Pi^{-1} \mathbf{X}_a \Pi = -\mathbf{X}_a$  for each  $\mathbf{X}_a$ . Thus  $\Pi^{-1} \mathbf{D} \Pi = -\mathbf{D}$ .

Now suppose that the initial and final states of the atom are eigenstates of  $\Pi$ , with eigenvalues  $\eta_i$  and  $\eta_f$ , respectively. Then<sup>38</sup>

$$\eta_i \eta_f \langle f | \mathbf{D} | i \rangle = \langle f | \Pi^{-1} \mathbf{D} \Pi | i \rangle = -\langle f | \mathbf{D} | i \rangle \quad (4.112)$$

---

<sup>38</sup>Note that since the parity operator  $\Pi$  is unitary its quantum numbers  $\eta$  behave multiplicatively, whilst those of a Hermitian generator of a continuous symmetry (such as  $\mathbf{P}$ ) behave additively.



**Figure 6:** Most atomic transitions are due to the absorption or emission of a photon from a single electron in the atom. The photon's angular frequency  $\omega = \Delta E/\hbar$ .

so the amplitude for the transition will vanish unless the initial and final atomic states have opposite parity,  $\eta_i \eta_f = -1$ . In the most common case, just a single electron is involved in the transition and in this case equation (4.110) shows that  $\eta_i \eta_f = (-1)^{\ell_i + \ell_f}$ , where  $\ell_{i,f}$  are the total orbital angular momentum quantum numbers of the initial and final states of this electron. Thus, in any radiative atomic transition involving just a single electron, the orbital angular momentum quantum number  $\ell$  must change by an odd number. If our initial and final states kept track of the photon as well as the atom, we'd be able to say overall parity is conserved in such transitions if we assign an intrinsic parity  $-1$  to the photon that is emitted or absorbed in the transition. (The intrinsic parity of the photon is related to the fact that the electromagnetic vector potential  $\mathbf{A}$ , whose quantum version provides the description of the photon, indeed transforms as a vector, not a pseudovector.) In fact,  $[H, \Pi] = 0$  holds in general for particles travelling in the presence of any type of electromagnetic interaction.

Having  $[H, \Pi]$  be zero means that if at time  $t = 0$  you set up a system to be a mirror image of another system, then their subsequent evolution will be identical in the sense that if you observe them at a later time, they will still be mirror images of one another. Hence, when  $[H, \Pi] = 0$  it is impossible to tell whether a system is being observed directly or through a mirror. One of the major surprises of twentieth-century physics was an experiment by Wu *et al.* in 1957, which showed that in fact the Hamiltonian of our Universe has  $[H, \Pi] \neq 0$  — you can see things in a mirror that are impossible in our World!