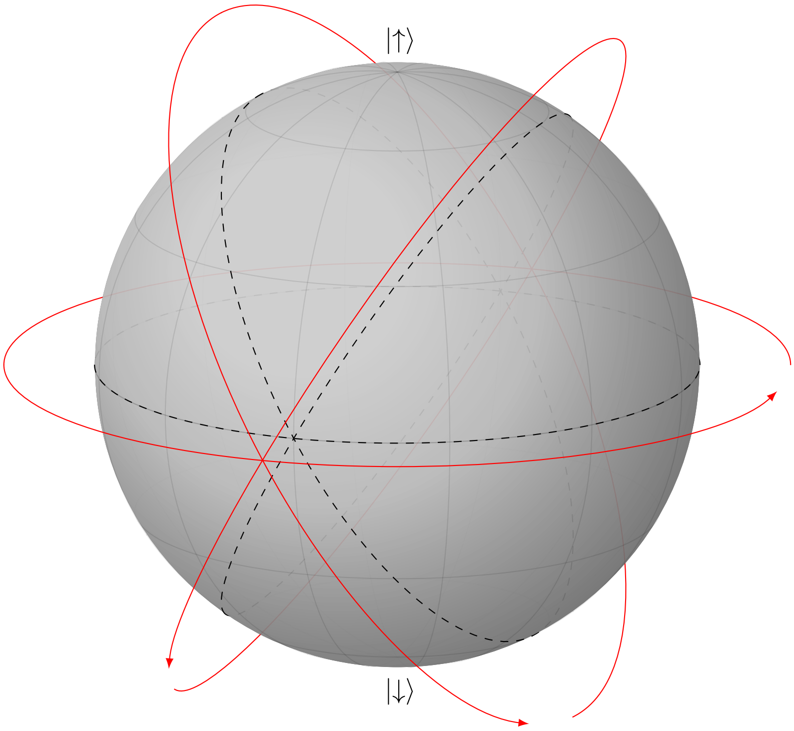
FURTHER QUANTUM THEORY

OXFORD UNIVERSITY MATHEMATICS

CHRISTOPHER BEEM
MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD



HILARY TERM 2026

[LAST UPDATED 2026-01-17]

Welcome to the 2025-2026 academic year incarnation of **B7.3 Further Quantum Theory** on the Oxford University Mathematics course. This course builds on the material from **A11 Quantum Theory**. In that first course, many core concepts of quantum theory (often called *quantum mechanics*, in reference to its status as a refinement of the classical mechanics of particles that you've studied in your Dynamics course, as well as possibly in **B7.1 Classical Mechanics**) were introduced. Here we will develop the subject more broadly and deeply, giving additional time and emphasis to the abstract mathematical formalism of the theory while also developing important technical methods that are used to analyse quantum mechanical systems in practice. Along the way, concepts from the theories of functional analysis, Lie groups, and representation theory will arise and be introduced. (Students are not expected to have much familiarity with these subjects in advance.) The rigorous foundations of quantum theory are thoroughly intertwined with deep aspects of functional analysis, and as a consequence our treatment of many topics will by necessity gloss over interesting subtleties. Whenever possible, I will indicate where a mathematically precise version of a statement exists but involves more machinery than we have at our disposal before giving a simplified treatment that suits our purposes.

Synopsis

- Abstract formulation of quantum mechanics in terms of linear operators on Hilbert spaces; Dirac notation; discrete and continuum states; time evolution and the propagator.
- Systems of several particles and Hilbert space tensor products; distinguishable and indistinguishable particles; Fermi–Dirac and Bose–Einstein statistics; the Pauli exclusion principle; elementary aspects of quantum entanglement.
- Symmetries in quantum mechanics as unitary and anti-unitary operators; rotations, angular momentum, and spin; spin-$\frac{1}{2}$ and projective representations of $\mathrm{SO}(3)$; addition of angular momentum; spin-statistics theorem; tensor operators and the Wigner–Eckart theorem.
- Approximation methods: Rayleigh–Schrödinger perturbation theory; variational methods; WKB approximation and Bohr–Sommerfeld quantisation.
- Elementary scattering theory in one dimension; relation between bound states and of poles/zeroes of the S-matrix.

References

You can never have enough quantum mechanics textbooks (I count around fifteen on the bookshelves of my office as I write this). Accordingly, you should feel free to utilise many sources and find some that are written in a way that you find appealing. This being said, the primary textbook references for this course are as follows:

- S. Weinberg, *Lectures on Quantum Mechanics* (CUP 2015).
- J. Sakurai and J. Napolitano, *Modern Quantum Mechanics* (CUP 2017).
- D. Griffiths and D. Schroeter, *Introduction to Quantum Mechanics* (CUP 2018).
- K. Hannabuss, *An Introduction to Quantum Theory* (OUP 1997).

However, we won't strictly adhere to the notational conventions of any of those. The following are also recommended for reference:

- E. Merzbacher, *Quantum Mechanics* (Wiley International 1970).
- P. A. M. Dirac, *The Principles of Quantum Mechanics* (OUP 2011).
- R. Shankar, *Principles of Quantum Mechanics* (Springer 1994).

The first is an older text that treats a number of topics in a bit more detail. In particular, the section on *Wave Packets and Free Particle Motion* in Merzbacher is nice, as well as the treatment of the WKB approximation. The second is a classic, written by the master himself. The third is a more-comprehensive-than-average text that has a lot of worked out examples, aimed more towards physicists.

Mathematical purists and rigour enthusiasts may want to consult the (quite advanced) textbooks:

- B. Hall, *Quantum Theory for Mathematicians* (Springer 2013).
- V. Moretti, *Fundamental Mathematical Structures of Quantum Theory* (Springer 2019).

These are both aimed well above the level of the present course, but (unlike most of the other references) they were really written for mathematicians rather than physicists.

Note on these notes

These notes are meant to complement and supplement the lectures for this course. They may be updated over the course of the term in response to developments in and feedback from the lectures and classes. Student feedback on the notes is very welcome, especially if there are typos or places where the text lacks clarity. Please send comments and corrections to `christopher.beem@maths.ox.ac.uk`.

Contents

Prologue

Chapter 1. Postulates and Examples

Chapter 2. Dirac's Formalism and Continuum Normalisation

Chapter 3. Composite Systems, Tensor Products, Entanglement

Chapter 4. Identical Particles and Statistics

Chapter 5. Symmetries and One-Parameter Unitary Groups

Chapter 6. Rotations, Angular Momentum, and Their Representations

Chapter 7. Intrinsic Spin and Addition of Angular Momentum

Chapter 8. Rayleigh–Schrödinger Perturbation Theory

Chapter 9. Variational Methods

Chapter 10. WKB Approximation

Chapter 11. One Dimensional Scattering

Chapter 12. Epilogue

Appendix A. Hilbert Space Minutiæ

Prologue

Review of Wave Mechanics and Examples

A first course on quantum theory traditionally focuses on the study of the Schrödinger equation (primarily time-independent, but occasionally time-dependent) for single-particle wave functions in one, two, and three spatial dimensions. These topics fall within the realm of what is often called *wave mechanics*. We include in this preparatory section a brief review of that formalism along with a recap of several standard examples that should be familiar from your previous studies. In the remainder of this course *we will freely refer back to these standard results*. This is not meant to be a comprehensive review of the content from **A11 Quantum Theory**; you should be prepared to refer back to material from that course when necessary.

WARNING: THIS SECTION WILL NOT BE THE SUBJECT OF ANY LECTURES.

0.1 Review of wave mechanics

The primary object of wave mechanics is Schrödinger's *wave function* for a particle moving in, say, one spatial dimension. This is a $\mathbb{C}$-valued function, often denoted Ψ, of real position and time variables x and t,

$$\Psi : \mathbb{R}_x \times \mathbb{R}_t \to \mathbb{C} \,. \tag{0.1}$$

The wave function is conventionally normalised to obey (at any time),

$$\int_{-\infty}^{\infty} \mathrm{d}x \, |\Psi(x,t)|^2 = 1 \,. \tag{0.2}$$

When so normalised, the wave function encodes (in particular) the probability density, ρ, for detecting the presence of the particle in question at a given point and at a given time according to

$$\rho(x,t) = |\Psi(x,t)|^2 \,. \tag{0.3}$$

The normalisation condition (0.2) ensures that with this interpretation in place, the total probability for finding the particle *anywhere at all* at any given time is one. Alternatively, one can work with wave functions that don't necessarily obey (0.2), in which case we have the modified expression for the probability density,[1]

$$\rho(x,t) = \frac{|\Psi(x,t)|^2}{\int_{-\infty}^{\infty} \mathrm{d}s \, |\Psi(s,t)|^2} \,. \tag{0.4}$$

If a particle moves in one spatial dimension subject to an external potential energy function $V(x)$, then the *time-dependent Schrödinger equation*—a partial differential equation that encodes time evolution of the wave function—takes the form

$$i\hbar \frac{\partial \Psi}{\partial t}(x,t) = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2}(x,t) + V(x)\Psi(x,t) \,. \tag{0.5}$$

Here $\hbar$ is the reduced Planck constant, which is a fundamental unit of angular momentum (*i.e.*, it has units of $[mass] \times [length]^2 \times [time]^{-1}$).[2] It is a simple matter of differentiating under the integral, using the time-dependent Schrödinger equation, and integrating by parts to confirm that if $\Psi(x,t)$ evolves in time according to (0.5) then the normalisation condition (0.2) will hold for all time t if it holds at any given time t_0.

---

[1] Here we demand that Ψ be squared-normalisable so the denominator is well-defined. This is the normal requirement for a wave function to be physically meaningful, though in the latter part of this course we will assign some meaning to certain non-normalisable wave functions.

[2] In SI units, the reduced Planck constant is $\hbar \approx 1.0546 \times 10^{-34}\, kg \cdot m/s^2$. In a quantum mechanical world, it is often a better idea to make a choice of units for which $\hbar = 1$ (so-called natural units). At a certain point in this course, it will make sense to think of $\hbar$ more like a formal small parameter.

The differential operator (with respect to x) that acts on the wave function on the right hand side of this equation is the *Hamiltonian operator* for the theory. If we introduce operators P and X that act on wave functions according to

$$(P\Psi)(x,t) = -i\hbar\frac{\partial\Psi}{\partial x}(x,t)~, \qquad (X\Psi)(x,t) = x\Psi(x,t)~, \tag{0.6}$$

then the Hamiltonian takes the form of the classical energy (alternatively, the Hamiltonian of the corresponding classical system, *cf.* **B7.1 Classical Mechanics**) with momentum and position replaced by their operator counterparts,

$$H = \frac{P^2}{2m} + V(X)~. \tag{0.7}$$

The time-dependent Schrödinger equation then takes the abstract form

$$i\hbar\frac{\partial\Psi}{\partial t}(x,t) = (H\Psi)(x,t)~. \tag{0.8}$$

In sufficiently well-behaved cases,[3] we can look for a basis of separable solutions to this equation of the form

$$\Psi(x,t) = \psi_n(x)\exp\left(-\frac{iE_nt}{\hbar}\right)~, \tag{0.9}$$

in which case the (real) numbers $\{E_n\}_{n=1,2,3...}$ are identified with the *energies* of the particle when described by the corresponding solution, and the functions $\{\psi_n\}$ are *stationary state wave functions* (depending only on position) that obey the *time-independent Schrödinger equation*,

$$(H\psi_n)(x) = -\frac{\hbar^2}{2m}\psi_n''(x) + V(x)\psi_n(x) = E_n\psi_n(x)~. \tag{0.10}$$

The time-dependence for a stationary state is determined by its energy through the time-dependent phase in (0.9). With the stationary state wave functions $\{\psi_n\}$ forming a basis for the space of possible wave functions at a fixed time, we can (in principle) understand time-evolution in complete generality,

$$\Psi(x,0) = \sum_n a_n\psi_n(x) \quad\Longrightarrow\quad \Psi(x,t) = \sum_n a_n\exp\left(-\frac{iE_nt}{\hbar}\right)\psi_n(x)~. \tag{0.11}$$

This follows from the linearity of the time-dependent Schrödinger equation and is one of the fundamental and unusual features of the quantum world.

0.2 Standard examples

When studying advanced methods for studying quantum mechanical systems, it is useful to be equipped with a small arsenal of standard examples from wave mechanics where the stationary-state wave functions ψ_n are understood exactly. In this subsection, we review several of the most important of these for ease of reference in the future, and also to fix certain notational conventions that will be recycled later in the course.

Example 0.2.1 (Particle in a box). When a particle is restricted to move in a fixed interval, say $x \in [0,a]$ for some positive real number a, the system is referred to as the particle in a (one-dimensional) box, or sometimes a particle in an infinite square well. As part of the definition of the system, one specifies that the wave function vanishes at the endpoints of the interval. (Alternatively, one could impose periodic boundary conditions and so describe a particle moving on a circle of circumference a.) This is sometimes phrased as a particle moving on the real line with a potential that is zero in the given interval and (positive) infinity outside of it.

---

[3] We will not dwell on the question of under which conditions this is a sound strategy; it will be the case for well-posed problems in quantum mechanics.

The stationary state wave functions for the particle in the box satisfy the simple boundary value problem,

$$-\frac{\hbar^2}{2m}\psi_n''(x) = E_n\psi_n(x)\;, \qquad \psi_n(0) = \psi_n(a) = 0\;, \tag{0.12}$$

This can be solved by inspection. The wave functions must be *sin* functions, and the boundary condition at a further constrains the possible energies E_n, giving the following set of normalisable solutions,

$$\psi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi x}{a}\right)\;, \qquad E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}\;, \qquad n = 1, 2, 3, \ldots\;. \tag{0.13}$$

The numerical prefactor is chosen so the wave functions will satisfy the standard normalisation condition; more generally we have the orthogonality relation

$$\int_0^a \mathrm{d}x\; \overline{\psi_n(x)}\psi_m(x) = \delta_{n,m}\;. \tag{0.14}$$

These functions form an orthonormal basis for the space of wave functions on the interval in the sense of Fourier series (so a general function gets represented as a suitably convergent infinite sum of multiples of basis functions). Time-evolution of any given wave function (taken as the wave function at time $t = 0$) is then determined by decomposing it according to its Fourier representation and introducing time-dependent phases depending on the energies in (0.13).

If one were to master only a single example in quantum mechanics, it should be the simple harmonic oscillator. The general structure of the quantum mechanical harmonic oscillator appears time and again in quantum physics, including, *e.g.*, in relativistic quantum field theory and string theory.

Example 0.2.2 (Harmonic Oscillator). Here a particle is allowed to range over the entire real line but with quadratic potential function $V(x) = \frac{1}{2}m\omega^2x^2$ (here by convention we take $\omega \in \mathbb{R}_+$), so the Hamiltonian operator is given by

$$H = \frac{1}{2m}P^2 + \frac{m\omega^2X^2}{2}\;. \tag{0.15}$$

The potential is parameterised such that the *classical* angular frequency of oscillation for the system would be ω.

The analysis of this system is simplified by introducing the first-order "ladder" operators,[4]

$$\alpha_\pm = \frac{1}{\sqrt{2m\hbar\omega}}(P \pm im\omega X)\;. \tag{0.16}$$

These obey several important relations,

$$[\alpha_-, \alpha_+] = 1\;, \tag{0.17}$$

$$[H, \alpha_\pm] = \hbar\omega\alpha_\pm\;, \tag{0.18}$$

$$H = \hbar\omega(\alpha_\pm\alpha_\mp \pm \tfrac{1}{2})\;. \tag{0.19}$$

Here we've adopted standard notation for the *commutator* of two differential operators A and B:

$$[A, B] := AB - BA\;. \tag{0.20}$$

Equation (0.18) implies that if $H\psi = E\psi$, then $H\alpha_\pm\psi = (E \pm \hbar\omega)\alpha_\pm\psi$. The $\alpha_\pm$ operators can therefore be understood to map between stationary states in the space of wave functions, shifting the energy by a definite amount.

Consequently, the ground state wave function ψ_0 (the stationary state wave function of least energy) must obey

$\alpha_-\psi_0 = 0$,[5] and this can be immediately solved to give the expression (with normalisation constant added)

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} \exp\left(-\frac{m\omega x^2}{2\hbar}\right) . \tag{0.21}$$

This stationary state has energy $E_0 = \frac{1}{2}\hbar\omega$ by virtue of (0.19), as can also be seen by direct calculation. All higher-energy states are obtained by the repeated action of α_+,

$$\psi_n = \frac{(-i\alpha_+)^n}{\sqrt{n!}}\psi_0 , \tag{0.22}$$

where the prefactor is such that these wavefunctions are all normalised and real. With a little bit of work, one finds that these stationary state wave functions can be expressed explicitly as follows,

$$\psi_n(x) = \frac{1}{\sqrt{2^n\, n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} \exp\left(-\frac{m\omega x^2}{2\hbar}\right) H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) , \tag{0.23}$$

where $H_n(x)$ is the n'th *Hermite polynomial*.[6] These form a basis for the space of normalisable wave functions on the real line, so the set of possible energies (the *spectrum*) of the quantum harmonic oscillator is given by

$$\left\{E_n = \hbar\omega\left(n+\frac{1}{2}\right) , \quad n = 0,1,2\ldots\right\} . \tag{0.24}$$

A (much) more involved, but still exactly solvable, example is the Hydrogen (or Hydrogen-like or Hydrogenic) atom. This is a problem of key historical importance; indeed much of the early development of Quantum Theory was motivated by the desire to understand the spectroscopic properties of atoms such as Hydrogen. A detailed analysis of the time-independent Schrödinger equation for this problem appeared in **A11 Quantum Theory**, including an introduction to angular momentum in quantum theory. We give a lightning review of the solution here. We will return to it several times throughout this course.

Example 0.2.3 (Hydrogenic atom). This is the problem of a single electrically charged particle (the electron) moving in three dimensions subject to a *Coulomb potential* due to a fixed-in-space pointlike nucleus[7] with positive electric charge Z times that of the electron in magnitude,[8]

$$V(\mathbf{r}) = -\frac{Zq_e}{|\mathbf{r}|} , \tag{0.26}$$

so the Hamiltonian operator is given by the following second-order differential operator when acting on wave functions $\psi(\mathbf{r})$,

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) . \tag{0.27}$$

The normalisable stationary states for this system are labelled by three *quantum numbers*: (n,ℓ,m). These are called the principal, orbital, and magnetic quantum numbers, respectively. Their ranges are restricted as

$$\begin{aligned} n &= 1,2,3,\ldots\ , \\ \ell &= 0,1,\ldots,n-1 , \\ m &= -\ell,-\ell+1,\ldots,\ell-1,\ell . \end{aligned} \tag{0.28}$$

---

[4] Beware, the conventions for the ladder operators here differ from some appearing in some text books by factors of $\pm i$.

[5] Though physically reasonable, the existence of a state with least energy requires an additional argument that we omit here. The argument appears in **A11 Quantum Theory**.

[6] The Hermite polynomials are defined as

$$H_n(x) = (-1)^n e^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(e^{-x^2}\right) . \tag{0.25}$$

The stationary state wave functions take the form,[9]

$$\psi_{n\ell m}(r,\theta,\varphi) = R_{n,\ell}(r) Y_\ell^m(\theta,\varphi) = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}\, R_{n,\ell}(r) P_\ell^m(\cos\theta) e^{im\varphi} \ . \tag{0.29}$$

The $Y_\ell^m(\theta,\varphi)$ are *spherical harmonic functions*, which obey

$$\mathbf{L}^2 Y_\ell^m(\theta,\varphi) = \hbar^2 \ell(\ell+1) Y_\ell^m(\theta,\varphi) \ , \qquad L_3 Y_\ell^m(\theta,\varphi) = \hbar m Y_\ell^m(\theta,\varphi) \ , \tag{0.30}$$

with respect to the *orbital angular momentum operators*,

$$\mathbf{L} = -i\hbar\,(\mathbf{x} \wedge \nabla) \ , \qquad L_3 = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) = -i\hbar\partial_\varphi \ . \tag{0.31}$$

The $P_\ell^m(\cos\theta)$ are the *associated Legendre polynomials*.[10] The radial wave function takes the form

$$R_{n,\ell}(r) = \exp\left(-\frac{Zr}{na_0}\right) f_{n,\ell}(r) \ , \tag{0.32}$$

where $f_{n,\ell}(r)$ is a polynomial of degree $n-1$ and $a_0 = \hbar^2/(mq_e^2)$ is the *Bohr radius*. The energy, E_n, depends only on the principal quantum number and is given by

$$E_n = -\frac{mq_e^4 Z^2}{2\hbar^2 n^2} = -\frac{q_e^2 Z^2}{2a_0 n^2} = -\frac{\alpha^2 mc^2 Z^2}{2n^2} \ , \tag{0.33}$$

where in the last expression, $\alpha = q_e^2/\hbar c \approx 1/137$ is the *fine structure constant*. Note that these stationary state energies are always negative, a fact which mirrors the negativity of the energy of bound orbits in the corresponding classical system.

In contrast to the previous two examples, the stationary state wave functions for the Hydrogen atom *do not* form a basis for all possible normalisable wave functions for the electron in a Coulomb potential. Rather, they form a basis only for the *bound state wave functions*.[11] The possibility of additional wave functions for the Hydrogen atom that are not bound states becomes especially relevant in the context of scattering theory.

---

[7] One need not take the nucleus to be fixed in space. Allowing for translational degrees of freedom for both electron and nucleus, one obtains this problem as that for the relative separation of the two with m the reduced mass of the system (and with overall center-of-mass degree of freedom decoupled).

[8] We adopt Gaussian units to avoid factors of $4\pi\varepsilon_0$. Here q_e is the charge of the electron, often written as e.

[9] The normalisation in the last term in (0.29) is so that

$$\int_{\mathbb{S}^2} Y_\ell^m(\theta,\phi) Y_{\ell'}^{m'}(\theta,\phi)\, \mathrm{d}\theta\, \mathrm{d}\phi = \delta_{\ell,\ell'}\delta_{m,m'} \ . \tag{0.34}$$

[10] We recall that the (misnamed) associated Legendre polynomials, which in general are not polynomials, are given for non-negative m by

$$P_\ell^m(x) = \frac{(-1)^m}{2^\ell \ell!}(1-x^2)^{\frac{m}{2}} \frac{\mathrm{d}^{\ell+m}}{\mathrm{d}x^{\ell+m}}(x^2-1)^\ell \ , \tag{0.35}$$

while for negative m they are determined by the relation

$$P_\ell^{-m}(x) = (-1)^m \frac{(\ell-m)!}{(\ell+m)!} P_\ell^m(x) \ . \tag{0.36}$$

Substituting $x = \cos\theta$, we see that the $P_\ell^m(\cos\theta)$ are indeed polynomials (of degree ℓ) in $\cos\theta$ and $\sin\theta$.

[11] We do not give a technical definition of bound state here, but physically it corresponds to states where the electron stays localised in the neighborhood of the nucleus instead of running off to infinity, analogous to the classical distinction between bound (elliptical) orbits versus parabolic or hyperbolic orbits in the Kepler problem.

Chapter 1

Postulates and Examples

In this chapter we introduce (or in some cases recall) the foundational postulates of quantum theory. You will have encountered these in a less detailed form in **A11 Quantum Theory**. Here we strive for accuracy, while stopping short of a completely rigorous discussion (which would require a good deal of background material from functional analysis). Where we gloss over technical subtleties there will often be a footnote or comment to this effect. Interested students are encouraged to take a look in the (advanced) textbooks by Hall and by Moretti for much more discussion.

After introducing the postulates, we will inspect two standard examples of quantum systems to highlight the general quantum theoretic framework. These examples will be the two-state *qubit* system and the system of a single particle moving on the real line.

1.1 Postulates of quantum theory

We adopt the abstract, algebraic formulation of quantum mechanics in terms of operators acting on Hilbert spaces. This is the main language of the subject, and was developed to a high degree of completeness by P. A. M. Dirac and John von Neumann; people often refer to these as the *Dirac–von Neumann axioms* of quantum theory. Many aspects of the original theory have come to be better understood mathematically since the inception of the framework, especially in connection with infinite-dimensional Hilbert spaces and their attendant subtleties. Here we will introduce, with comments, a set of working postulates/axioms for how physical systems should be described within the context of quantum theory. A thorough familiarity with abstract linear algebra and metric/inner product spaces is assumed.

Postulate I (Space of states). *States of a physical system correspond to rays in a complex Hilbert space.*

This postulate requires some definitions, which we reproduce here. (Appendix A provides a more detailed review of the definition of a Hilbert space and some discussion of advanced topics for those who are curious.)

Definition 1.1.1 (Hilbert Space). A complex Hilbert space is a complex vector space $\mathcal{H}$ (possibly infinite-dimensional) with an Hermitian inner product $(\cdot,\cdot):\mathcal{H}\times\mathcal{H}\to\mathbb{C}$.

An Hermitian inner product is, in particular, a positive definite *sesquilinear form*, so for any $\varphi_i,\psi_j\in\mathcal{H}$, and $\alpha,\beta\in\mathbb{C}$ we have,[12]

$$\begin{aligned}(\varphi_1+\varphi_2,\psi_1+\psi_2) &= (\varphi_1,\psi_1)+(\varphi_1,\psi_2)+(\varphi_2,\psi_1)+(\varphi_2,\psi_2)\,,\\ (\alpha\varphi,\beta\psi) &= \overline{\alpha}\beta(\varphi,\psi)\,,\end{aligned}\tag{1.1}$$

Hermiticity amounts to the additional relation

$$(\varphi,\psi)=\overline{(\psi,\varphi)}\,.\tag{1.2}$$

For $\mathcal{H}$ infinite dimensional, it is also required that $\mathcal{H}$ be complete (limits of Cauchy sequences must exist).

Furthermore, it is normally assumed that in the infinite dimensional case the Hilbert spaces dealt with in quantum theory are *separable*, meaning they admit a countable basis. The issues of completeness and separability will not play any significant role in this course, though they are important for providing rigorous foundations to the subject.

Definition 1.1.2 (Ray). A *ray* in $\mathcal{H}$ is the set of non-zero, scalar multiples of a given non-zero vector.

---

[12] Observe that we adopt "physics conventions", in which the inner product is conjugate-linear in the *first* argument. In the (non-physical) mathematical literature, it is more common to have the second entry be conjugate linear. This is clearly a matter of convention.

Alternatively, let us introduce an equivalence relation $\sim$ on the set of non-zero vectors in $\mathcal{H}$,

$$\psi \sim \alpha\psi \qquad \forall\, \alpha \in \mathbb{C}^{\times} \,. \tag{1.3}$$

Then a ray is an equivalence class with respect to $\sim$. This characterisation of rays in $\mathcal{H}$ lets us identify the *space of quantum states* with the quotient of the space of nonzero vectors $\mathcal{H} \setminus \{0\}$ by the equivalence relation, *i.e.*,

$$\{\text{Quantum States}\} \;=\; \left({}^{\mathcal{H}\setminus\{0\}}\big/_{\sim}\right) \;\cong\; \mathbb{P}(\mathcal{H})\,. \tag{1.4}$$

The latter equivalence identifies this space with the *projectivisation* of the Hilbert space (*cf.* **ASO Projective Geometry**). It is often convenient to adopt the convention of working with *normalised* state vectors (as in the discussion of wave functions in Chapter 0. This leaves an overall phase ambiguity, so we also have

$$\mathbb{P}(\mathcal{H}) \cong \left({}^{\{\psi\in\mathcal{H}\,|\,(\psi,\psi)=1\}}\big/_{(\psi\sim e^{i\phi}\psi\,,\ \phi\in\mathbb{R})}\right)\,. \tag{1.5}$$

The two characterisations are completely equivalent, but by working with normalised vectors one often ends up with superficially simpler formulæ.

Remark 1.1.3. The interplay between the physical space of states, $\mathbb{P}(\mathcal{H})$, and the larger Hilbert space $\mathcal{H}$ underlies several interesting phenomena in quantum theory. By working in Hilbert space, which in particular is a vector space, one makes manifest the important linear aspects of quantum theory. However, as we shall see in our discussion of symmetries and, in particular, of spin, it is sometimes important not lose sight of the true space of states being the projectivised Hilbert space.

Postulate II (Observables)**.** *Observables of a physical system correspond to self-adjoint operators on the Hilbert space $\mathcal{H}$.*

Recall that a linear map $A : \mathcal{H} \to \mathcal{H}$ is referred to in the quantum mechanical setting (amongst other places) as an *operator* (as in differential operator).

Definition 1.1.4 (Adjoint)**.** The adjoint of a linear operator $A : \mathcal{H} \to \mathcal{H}$ is another operator $A^* : \mathcal{H} \to \mathcal{H}$ such that for any $\varphi, \psi \in \mathcal{H}$ we have,

$$(\varphi, A\psi) = (A^*\varphi, \psi)\,. \tag{1.6}$$

For the case of finite-dimensional $\mathcal{H}$ (identified with $\mathbb{C}^n$ with the standard inner product after making a choice of orthonormal basis), operators are just $n \times n$ complex matrices, and the adjoint is the transpose-conjugate. There is a complication in the case of infinite-dimensional $\mathcal{H}$. There, linear operators are generally only partially defined, so their domain $D(A) \subsetneq \mathcal{H}$. (These are referred to as *unbounded operators*.) The adjoint of an unbounded operator then has its own domain $D(A^*)$, and self-adjointness requires $D(A) = D(A^*)$ which is not automatic. In this course we will not generally discuss the domains of the observables we study in any systematic way. See Appendix A for a bit more advanced discussion of this.

There is an important and, in the general case, deep result that we will utilise (often implicitly) throughout this course: the *spectral theorem for self-adjoint operators on a Hilbert space*. The statement of this theorem in the general case already involves a lot of technology. *Roughly speaking*, the theorem says that a self-adjoint operator on a Hilbert space admits a complete (orthonormal) basis of eigenvectors, so for a general observable we can write something like,

$$\psi = \sum_n c_n \psi_n\,, \tag{1.7}$$

where the ψ_n are eigenstates of the observable A,

$$A\psi_n = a_n \psi_n\,. \tag{1.8}$$

The set of eigenvalues $\{a_n\}$ is then referred to as the *spectrum* of the operator A.

Remark 1.1.5. The rough characterisation of the spectral theorem above omits an important subtlety associated with infinite-dimensional Hilbert spaces, which is the possibility of a *continuous spectrum*. We will return to this in the next chapter. For now you should take the above characterisation as impressionistic in general, but accurate in many important examples.

Postulate III (Measurement). *When measuring an observable A, the only possible results correspond to the elements of the spectrum of A. The probability of a given result a is the squared norm of the orthogonal projection of the (normalised) initial state onto the a-eigenspace $\mathcal{H}_a \subseteq \mathcal{H}$.*

In the case when all eigenvalues of A are nondegenerate, this means that for a (normalised) state ψ as in (1.7), the probability of obtaining, say, a_i when measuring A is exactly $|c_i|^2$. More generally, suppose that for some subset $\{\psi_{i\in I}\}$ of the A eigenbasis, the A eigenvalues are all degenerate $a_{i\in I} = a$. Let Π_a denote the orthogonal projection operator onto the a eigenspace $\mathcal{H}_a \subset \mathcal{H}$ for which the $\{\psi_{i\in I}\}$ form a basis, *i.e.*, the operator that acts on a state as given in (1.7) according to

$$\Pi_a \psi = \sum_{i\in I} c_i \psi_i \,. \tag{1.9}$$

Then the probability of observing a when measuring A is given by

$$(\Pi_a\psi, \Pi_a\psi) = \left(\sum_{i\in I} c_i\psi_i, \sum_{i\in I} c_i\psi_i\right) = \sum_{i\in I} |c_i|^2 \,. \tag{1.10}$$

These probabilistic statements are compatible with the following definitions.

Definition 1.1.6 (Expectation Value). The *expectation value* of the observable A in a state ψ is given by

$$\mathbb{E}_\psi(A) \equiv \langle A\rangle_\psi = (\psi, A\psi) = \sum_n a_n |c_n|^2 \,. \tag{1.11}$$

Definition 1.1.7 (Dispersion). The *dispersion* of the observable A in the state ψ is given by

$$\Delta_\psi(A) = \mathbb{E}_\psi\left(\left(A - \langle A\rangle_\psi\right)^2\right) = \mathbb{E}_\psi\left(A^2 - \langle A\rangle_\psi^2\right) \,. \tag{1.12}$$

These definitions agree with the usual statistical notion of the expectation and variance of a random variable.[13]

Postulate IV (Wave function collapse). *Immediately following a measurement of the observable A that yields the result a, the state of the system will be the orthogonal projection of the initial state onto the a-eigenspace.*

This postulate is the subject of quite a lot of discussion—this usually takes place under the banners of *interpretations of quantum mechanics* and *the measurement problem*. These discussions sometimes have a philosophical flavour and will not be pursued in this course; indeed we will spend very little or no time discussing measurement. Some of the (more technical than interpretative) aspects of quantum measurement play an important role in quantum information theory, and if you're interested you might look into **C7.4 Introduction to Quantum Information**.

As a practical matter, the statement of wave function collapse has the reasonable consequence that if one measures an observable A and finds some value a, then *immediately* measuring A again will reproduce the result a with absolute certainty.

---

[13] In the above discussion, our formulæ were all tailored to the case of a normalised state vector ψ. For general ψ, one must divide by the normalising factor (ψ, ψ) in both cases.

Remark 1.1.8 (Compatible and Incompatible Measurements). If two observables A and B commute, so $[A, B] = 0$,[14] then there exists (in the same sense as in Postulate II) a basis of states which are simultaneous eigenstates of A and B. In this case, one can unambiguously observe A and B simultaneously because the projection operators onto the appropriate eigenspaces commute. On the other hand, if $[A, B] \neq 0$ then measurement of A and B are incompatible, in the sense that if one measures A it will effect the outcome of a measurement of B and *vice versa*. In the special case of position and momentum operators, this idea is encapsulated in the Heisenberg uncertainty relation that you have seen in your previous course.

Postulate V (Time evolution). *The time development of a given state ψ is controlled by a special observable called the Hamiltonian, usually denoted by H, according to the general time-dependent Schrödinger equation,*

$$i\hbar \frac{\mathrm{d}\psi}{\mathrm{d}t} = H\psi \,. \tag{1.13}$$

In general, H as an operator can depend explicitly on time $H = H(t)$, but this doesn't effect the form of the time-dependent Schrödinger equation. (In most of the examples we study this will not be the case; we study time-independent Hamiltonians.)

An important consequence of this equation is that the inner product (a.k.a. the *overlap*) between any two state vectors is preserved under time evolution,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(\varphi, \psi) &= \left(\frac{\mathrm{d}\varphi}{\mathrm{d}t}, \psi\right) + \left(\varphi, \frac{\mathrm{d}\psi}{\mathrm{d}t}\right) \,, \\ &= \left(\frac{H\varphi}{i\hbar}, \psi\right) + \left(\varphi, \frac{H\psi}{i\hbar}\right) \,, \\ &= \frac{i}{\hbar}(H\varphi, \psi) - \frac{i}{\hbar}(\varphi, H\psi) \,, \\ &= 0 \,, \end{aligned} \tag{1.14}$$

where we have used the self-adjointness of H. In particular, setting $\varphi = \psi$, this implies that normalised state vectors remain normalised under time evolution.

We can say a bit more about how to formalise time evolution for *finite intervals of time* as follows. (We will also revisit some of this material in a later section.) Let us define an operator $U(t_1; t_0) : \mathcal{H} \to \mathcal{H}$ that sends a state vector (thought of as the state of our system at time t_0), call it ψ_{t_0}, to its time evolution forward to time $t_1 > t_0$, which we call ψ_{t_1}. Because (1.13) is linear, then $U(t_1; t_0)$ is itself a linear operator. We now have

$$\begin{aligned} (\varphi_{t_1}, \psi_{t_1}) &= (U(t_1; t_0)\varphi_{t_0}, U(t_1; t_0)\psi_{t_0}) \,, \\ &= (\varphi_{t_0}, U(t_1; t_0)^* U(t_1; t_0)\psi_{t_0}) \,, \\ &= (\varphi_{t_0}, \psi_{t_0}) \,. \end{aligned} \tag{1.15}$$

with the last equality a consequence of the time-independence of overlaps. We conclude that the time-evolution operator $U(t_1; t_0)$ obeys the important relation

$$U(t_1, t_0)^* = U(t_1, t_0)^{-1} \,. \tag{1.16}$$

Such operators are called *unitary operators*.

---

[14] Once again, there is a small subtlety with defining this commuting property for general operators in infinite-dimensional Hilbert spaces, but we won't encounter any examples where the subtlety is relevant

Definition 1.1.9 (Unitary Operator)**.** A *unitary operator* U on a Hilbert space is a linear map $\mathsf{U} : \mathcal{H} \to \mathcal{H}$ that obeys

$$\mathsf{U}^* \mathsf{U} = \mathsf{U} \mathsf{U}^* = 1_{\mathcal{H}} . \tag{1.17}$$

This is, equivalently, a surjective map from $\mathcal{H}$ to $\mathcal{H}$ obeying $\mathsf{U}^*\mathsf{U} = 1_{\mathcal{H}}$.

The requirement to separately consider left- and right-composition by U^* or to demand surjectivity is associated with the subtleties of infinite-dimensional Hilbert space. In finite dimensional settings, and for a given choice of orthonormal basis, a unitary operator is just the same thing as a unitary matrix.

When the Hamiltonian itself is time-independent, the time evolution operator will only depend on the time interval $t_1 - t_0$ and we can write $U(t_1 - t_0)$ instead, or just $U(t)$ for simplicity. To understand time-evolution then amounts to identifying the basis of states that diagonalises the action of the Hamiltonian, *i.e.*, the stationary states obeying

$$H\psi_n = E_n \psi_n . \tag{1.18}$$

The time evolution operator $U(t)$ can be understood very simply in the basis of stationary states, with its action given by

$$U(t)\psi_n(x) = \exp\left(-\frac{iE_n t}{\hbar}\right) \psi_n(x) , \tag{1.19}$$

from which the action on a general state can be deduced by linearity. We observe that we can write this operator as an exponentiation of the Hamiltonian operator,

$$U(t) = \exp\left(-\frac{iHt}{\hbar}\right) , \tag{1.20}$$

where the expression on the right has an obvious interpretation when applied to stationary states, and the more general case follows by linearity.

1.2 Example: the qubit

The simplest instances of quantum systems are those with *finite-dimensional Hilbert spaces*. The simplest non-trivial example is when the Hilbert space has dimension two, in which case the system is a *qubit*.[15] Choosing any orthonormal basis we get a (non-canonical) identification $\mathcal{H} \cong \mathbb{C}^2$ with inner product

$$(\mathbf{u}, \mathbf{v}) = \overline{\mathbf{u}} \cdot \mathbf{v} . \tag{1.21}$$

The space of quantum states is the complex projective line (*i.e.*, the Riemann sphere), $\mathbb{P}(\mathbb{C}^2) = \mathbb{CP}^1$. Topologically $\mathbb{CP}^1 \cong \mathbb{S}^2$, the two-sphere, so the space of quantum states of the qubit system is actually a sphere—this is sometimes called the *Bloch sphere*.

Observables in this qubit system are 2×2 self-adjoint (a.k.a. Hermitian) matrices. A basis for these is as follows,

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.22}$$

Here σ_0 is the identity operator, so does not play much of a role as an observable. The σ_i, $i = 1, 2, 3$ are the so-called *Pauli matrices*, which enjoy the commutation relations

$$[\sigma_1, \sigma_2] = 2i\sigma_3 , \qquad [\sigma_2, \sigma_3] = 2i\sigma_1 , \qquad [\sigma_3, \sigma_1] = 2i\sigma_2 . \tag{1.23}$$

In the given basis, σ_3 is diagonalised with eigenvalues ± 1. Meanwhile σ_1 and σ_2 are not diagonalised, and the non-trivial commutators imply that these are incompatible observables. Indeed, the σ_1 eigenstates are the vectors $(1, 1)^t$

---

[15] Qubits are the building blocks of quantum computers. They stand in as the quantum mechanical analogue of a classical bit, which is a degree of freedom that takes one of two values 0 and 1. Unlike classical bits, qubits can exist in any complex linear superposition of their two basis states, which accounts for their more powerful computational properties.

and $(1,-1)^t$, and those of σ_2 are $(1,-i)^t$ and $(1,i)^t$. By choosing one of these as a basis one could diagonalise σ_1 or σ_2 instead, or in fact an arbitrary real linear combination of the σ_i. We will encounter more properties of these matrices when we revisit this Hilbert space in the context of our treatment of rotations and spin.

For a qubit, the possible time-evolution operators are elements of the two-dimensional unitary group $\mathrm{U}(2)$ of 2×2 matrices whose transpose-conjugate (adjoint) is their inverse. It is interesting to consider the action of this group on the space of quantum states. In particular, note that there is a subgroup $\mathrm{U}(1)\subset\mathrm{U}(2)$ of matrices of the form

$$U=\begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{i\phi}\end{pmatrix} , \tag{1.24}$$

whose action on the space of quantum states is actually trivial (because states related by an overall phase are equivalent). Thus, we have that the action of $\mathrm{U}(2)$ on the space $\mathbb{CP}^1$ factors through the quotient by this $\mathrm{U}(1)$ subgroup, which happens to be isomorphic to the three-dimensional (special) orthogonal group,

$$\mathrm{U}(2)/\mathrm{U}(1)\cong\mathrm{SO}(3) . \tag{1.25}$$

There is an obvious action of $\mathrm{SO}(3)$ on $\mathbb{CP}^1$ which is just the rigid rotations of the two-sphere, and indeed this is how this action arises (we won't prove it here). We are most familiar with $\mathrm{SO}(3)$ in connection with rotations in three-dimensional space, and its appearance here is no accident, as we will see in more detail later in Chapter 6.

1.3 Example: Particle on the real line

The principal instance of this formalism treated in **A11 Quantum Theory** arose in describing the movement of a single point-particle in $d=1$, 2, or 3 dimensions. For now we restrict to $d=1$. In this case, a state vector is represented by a Schrödinger wave function $\psi:\mathbb{R}\to\mathbb{C}$, a complex-valued function of position $x\in\mathbb{R}$. The Hermitian inner product of two state vectors φ and ψ is given by

$$(\varphi,\psi)=\int_{\mathbb{R}}\overline{\varphi(x)}\psi(x)\,\mathrm{d}x . \tag{1.26}$$

Physical wave functions are required to be *square-normalisable* (sometimes we will just say "normalisable"):

$$(\psi,\psi)=\int_{\mathbb{R}}|\psi(x)|^2\,\mathrm{d}x<\infty . \tag{1.27}$$

The Hilbert space assigned to such a system is then, roughly speaking, the vector space of complex, square-integrable functions on $\mathbb{R}^d$. With some technical improvements,[16] this leads to the definition of the Hilbert space $\mathcal{H}\cong L^2(\mathbb{R})$ (here L is for Lebesgue and the exponent indicates that it is the absolute value *squared* that appears in the norm).

The observables that are most frequently discussed in this setting are realised as differential operators on wave functions. In particular, the most natural observables from a physical viewpoint are the momentum and position operators P and X, which act according to

$$(P\psi)(x)=-i\hbar\left(\frac{\mathrm{d}\psi}{\mathrm{d}x}\right)(x) , \qquad (X\psi)(x)=x\psi(x) . \tag{1.28}$$

More generally, one constructs many self-adjoint differential operators as observables by composing the actions of P and X appropriately.[17] For example, the Hamiltonian operator is usually taken to have its classical form

$$H=\frac{P^2}{2m}+V(X) , \tag{1.29}$$

---

[16] As you may know if you have taken **A4 Integration** or **B4.1 Functional Analysis**, to really define this space well one needs to form equivalence classes of functions that agree almost everywhere, *i.e.*, outside of sets of Lebesgue measure zero).

[17] However, recall that because in general $(AB)^*=B^*A^*$, a generic composition of P's and X's will not be self-adjoint despite P and X themselves being so.

for some function V that is frequently (though not always) bounded below. This acts on states-as-wave-functions according to

$$(H\psi)(x) = -\frac{\hbar^2}{2m}\psi''(x) + V(x)\psi(x)\ . \tag{1.30}$$

The stationary states, or H eigenstates, are those $\{\psi_n\}$, that obey the wave-function version of the time-independent Schrödinger equation (see the review in Chapter 0),

$$-\frac{\hbar^2}{2m}\psi_n''(x) + V(x)\psi_n(x) = E_n\psi_n(x)\ , \tag{1.31}$$

for E_n in the spectrum of H. For examples like the harmonic oscillator, these form an orthonormal basis for $L^2(\mathbb{R})$.

Remark 1.3.1. Though X and P are intuitively two very natural observables for this system, note that the space of all square-integrable functions includes many functions which are not differentiable (or even continuous!) and many functions which, after multiplication by x, would no longer be square normalisable. This is an instance of the aforementioned complication that observables are often only defined on a (dense) subspace of the Hilbert space in the infinite-dimensional case. X and P also don't admit proper eigenfunctions in $\mathcal{H}$, an issue that we will return to in the coming Chapter 2.

A class of observables that are quite well behaved, and which you in fact studied a bit in **A11 Quantum Theory** in different terms, are the projection operators

$$\begin{aligned} \Pi_E :\ \mathcal{H} &\longrightarrow \mathcal{H}\ , \\ \psi(x) &\longmapsto 1_E(x)\psi(x)\ , \end{aligned} \tag{1.32}$$

where $1_E(x)$ is the indicator function for a measurable set $E \subseteq \mathbb{R}$. This operator is easily verified to be self-adjoint and a projection (in that $\Pi_E \circ \Pi_E = \Pi_E$), which implies that its eigenvalues can only be zero or one. Indeed, by inspection one observes that a function can only be an eigenfunction if it is supported entirely within E (in which case it has eigenvalue one) or if it is supported entirely outside of E (in which case it has eigenvalue zero).

A measurement of Π_E corresponds to asking the yes/no question "is the particle located within the region E?". The expectation value for this operator is given by

$$\mathbb{E}_\psi(\Pi_E) = \int_{-\infty}^{\infty} \mathrm{d}x\,|\psi(x)|^2 1_E(x) = \int_E \mathrm{d}x\,|\psi(x)|^2\ , \tag{1.33}$$

which is exactly how in **A11 Quantum Theory** you computed the probability that the particle was located in the region E. Here we see that this result is reproduced in a framework where one only asks questions of a system that can be encoded in an observable.

Remark 1.3.2. The identification $E \to \Pi_E$ is an example of what is called a *projection-valued measure*, and this is the object whose existence is guaranteed by the application of (one formulation of) the spectral theorem to the position operator X.

Chapter 2

Dirac's Formalism and Continuum Normalisation

The subject of this section is, in some sense, a matter of notational formalism. We will (re-)introduce the *bra-ket* formalism of Dirac for representing states and observables in quantum mechanical systems. In reality, Dirac's formalism (in the broadest sense) is more substantive than just a change of notation. The novelty arises when discussing observables with a *continuous spectrum*. This can happen only in infinite-dimensional Hilbert spaces.

2.1 States, dual states, and matrix elements

The basic notational device introduced by Dirac is the *bra-ket*. Here we represent vectors in a Hilbert space as *kets*,

$$\psi \in \mathcal{H} \quad \longleftrightarrow \quad |\psi\rangle \ . \tag{2.1}$$

Since a Hilbert space is equipped with an inner product, we can also assign to a given state a dual vector

$$\begin{aligned} \varphi_\psi \ &: \ \mathcal{H} \to \mathbb{C} \,, \\ &: \ \chi \mapsto (\psi, \chi) \,. \end{aligned} \tag{2.2}$$

By sesquilinearity of the inner product on $\mathcal{H}$, this map is $\mathbb{C}$-conjugate-linear: $\varphi_{\lambda\psi} = \bar{\lambda}\varphi_\psi$ for $\lambda \in \mathbb{C}$. An important result in functional analysis is the following.

Theorem 2.1.1 (Riesz–Fréchet representation theorem)**.** Let $\mathcal{H}$ be a Hilbert space. For every continuous linear functional $\varphi \in \mathcal{H}^*$, there exists a unique $\psi \in \mathcal{H}$ such that $\varphi = \varphi_\psi$ (using the notations above).

We do not provide a proof here. In the finite-dimensional setting, it is not a difficult result, but in the infinite-dimensional case it is not as obvious. Indeed, the presence of the adjective *continuous* in the above theorem statement is relevant precisely in the case of infinite-dimensional $\mathcal{H}$ (all linear functionals are continuous in a finite dimensional Hilbert space). This theorem establishes a canonical *bijective, antilinear isometry* between $\mathcal{H}$ and $\mathcal{H}^*$.

In Dirac notation, we denote elements of the (continuous) dual space $\mathcal{H}^*$ by *bras*:

$$\varphi \in \mathcal{H}^* \quad \longleftrightarrow \quad \langle\varphi| \ . \tag{2.3}$$

As a somewhat overloaded notational convention, we often use as the label for a *bra* the name of the state in $\mathcal{H}$ to which it corresponds under the Riesz–Fréchet isometry,

$$\varphi_\psi \quad \longleftrightarrow \quad \langle\psi| \ . \tag{2.4}$$

With these conventions in place, we denote the inner product between two states φ and ψ as a composite *bra-ket*, where the state and dual state are fused together in the visually natural manner,

$$(\varphi, \psi) \quad \longleftrightarrow \quad \langle\varphi|\psi\rangle \ . \tag{2.5}$$

An operator A can act on states/kets from the left, whereupon we will notationally allow it to be "absorbed" into the ket

$$A\,|\psi\rangle = |A\psi\rangle \ . \tag{2.6}$$

Similarly, operators act from the right upon bras, and are replaced by their adjoint when absorbed into the bra,

$$\langle\varphi|\,A = \langle A^*\varphi| \ . \tag{2.7}$$

Within an inner product, an operator can be moved about accordingly,

$$(\varphi, A\psi) = \langle\varphi|A\psi\rangle = \langle\varphi|A|\psi\rangle = \langle A^*\varphi|\psi\rangle = (A^*\varphi, \psi)\ . \tag{2.8}$$

The quantity $\langle\varphi|A|\psi\rangle$ will be referred to as the *matrix element* of A between φ and ψ.

2.2 Constructions with bra-kets

The bra-ket formalism is convenient for representing an array of natural constructions involving linear operators on Hilbert spaces. For example, given the state $|\alpha\rangle$ and the dual-state $\langle\beta|$, we can construct the *outer product*,

$$\begin{aligned} |\alpha\rangle\langle\beta| \ :\ \mathcal{H} &\to \mathcal{H}\ , \\ |\psi\rangle &\mapsto |\alpha\rangle\langle\beta|\psi\rangle = (\langle\beta|\psi\rangle)\,|\alpha\rangle\ . \end{aligned} \tag{2.9}$$

So we can concatenate *bras* and *kets* in the visually obvious manner and get meaningful operations. Outer products also behave well under taking adjoints,[18]

$$(|\alpha\rangle\langle\beta|)^* = |\beta\rangle\langle\alpha|\ . \tag{2.10}$$

Now let $\{|i\rangle\ ,\ i\in I\}$ be an orthonormal basis for $\mathcal{H}$ (here I is some finite or countably infinite indexing set). Orthonormality means we have $\langle i|j\rangle = \langle j|i\rangle = \delta_{ij}$. We can then write an arbitrary vector in $\mathcal{H}$ uniquely as a (possibly infinite) linear combination of these basis vectors,

$$|\psi\rangle = \sum_{i\in I} c_i\,|i\rangle\ . \tag{2.11}$$

The components c_j for some $j\in I$ are extracted by acting with the *bra* corresponding to $|j\rangle$,

$$\langle j|\psi\rangle = \sum_{i\in I} c_i\,\langle j|i\rangle = \sum_{i\in I} c_i\delta_{ij} = c_j\ . \tag{2.12}$$

We see that we can realise the orthogonal projection Π_j onto the one-dimensional subspace spanned by the basis vector $|j\rangle$ using the outer product, $|j\rangle\langle j|$,[19]

$$|j\rangle\langle j|\psi\rangle = c_j\,|j\rangle\ . \tag{2.14}$$

More generally, for a linear subspace $\mathcal{H}'\subseteq\mathcal{H}$ with orthonormal basis $|i'\rangle\ ,i'\in I'$, we can form the manifestly self-adjoint, orthogonal projection operator from $\mathcal{H}$ onto $\mathcal{H}'$:

$$\Pi_{\mathcal{H}'} = \sum_{i'\in I'} |i'\rangle\langle i'|\ . \tag{2.15}$$

In particular, for the case $\mathcal{H}'=\mathcal{H}$, we have an expression for the identity operator,

$$\Pi_{\mathcal{H}} \equiv 1_{\mathcal{H}} = \sum_{i\in I} |i\rangle\langle i|\ . \tag{2.16}$$

This expression is often referred to as a *resolution of the identity* or *completeness relation*. Given a linear operator $A:\mathcal{H}\to\mathcal{H}$, we can then resolve it in terms of its matrix elements with respect to the given basis,

$$A = 1_{\mathcal{H}}A1_{\mathcal{H}} = \sum_{i,j\in I} |i\rangle\langle i|A|j\rangle\langle j| = \sum_{i,j\in I} A_{ij}\,|i\rangle\langle j|\ . \tag{2.17}$$

where

$$A_{ij} = \langle i|A|j\rangle\ . \tag{2.18}$$

---

[18] Verify this relation if it isn't obvious to you by inspection.

[19] For the a general state vector ψ, not necessarily normalised, we have the orthogonal projection operator,

$$\Pi_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}\ . \tag{2.13}$$

Finally, for A an observable if the states $\{|i\rangle\}$ are an orthonormal basis of A eigenstates obeying $A\,|i\rangle = a_i\,|i\rangle$ then we have matrix elements $A_{ij} = a_i\delta_{ij}$ and (2.17) becomes the *spectral decomposition* of A,

$$A = \sum_i a_i\,|i\rangle\,\langle i|\ . \tag{2.19}$$

Remark 2.2.1. In the case where $\mathcal{H}$ is finite-dimensional, this is all pretty familiar. The outer product $|i\rangle\,\langle j|$ corresponds to the matrix that is all zeroes except for having a one in the i'th row at the j'th column, and (2.17) describes the building up the operator A entry by entry as a matrix, while (2.19) corresponds to the matrix expression for A in the basis where A is diagonalised, which is the usual spectral decomposition of an Hermitian matrix. The resolution of the identity is just the expression for the identity operator as the identity matrix.

In terms of bra-kets, we can represent the expectation value of an observable as follows. If our basis $\{|i\rangle\}$ diagonalises the observable A as above, then we have

$$\begin{aligned}\mathbb{E}_\psi(A) = \mathbb{E}_\psi(A\mathbb{1}_{\mathcal{H}}) &= \sum_{i\in I}\langle\psi|A|i\rangle\,\langle i|\psi\rangle\ ,\\ &= \sum_{i\in I} a_i|\,\langle i|\psi\rangle\,|^2\ ,\end{aligned} \tag{2.20}$$

which matches the notion of expectation value for a random variable.

2.3 Continuous observables

We now come to the important issue of observables with *continuous spectrum*. In finite dimensional Hilbert spaces (and, it turns out, for something called a *compact operator* on an infinite-dimensional Hilbert space) the spectrum of any observable is discrete, being just the set of eigenvalues. For more general operators in infinite-dimensional Hilbert spaces we may potentially encounter a subtlety.

Just the definition of the spectrum of an operator is in fact more subtle in the infinite-dimensional case than just the eigenvalues. Indeed, we have the following:

Definition 2.3.1 (Spectrum). The *spectrum* of a self-adjoint operator A on a Hilbert space $\mathcal{H}$ is the subset $\sigma(A) \subseteq \mathbb{R}$ such that for $\lambda \in \sigma(A)$, the shifted operator $A - \lambda\mathbb{1}_{\mathcal{H}}$ does not have a (bounded, everywhere defined) inverse.

We will not dwell on the bounded/everywhere-defined caveats, which are relevant for a fully rigorous treatment. In finite dimensions, the equivalence of non-invertibility and λ being an eigenvalue is automatic upon consideration of the characteristic polynomial.

In the infinite-dimensional case, more elaborate situations are possible, and in particular the spectrum can include a continuum. Indeed, we can think of the position operator X acting on $L^2(\mathbb{R})$. Considering our definition, a value $\lambda \in \sigma(X)$ if for any $g \in L^2(\mathbb{R})$ we can't find an $f \in L^2(\mathbb{R})$ that solves the problem,

$$(x-\lambda)f(x) = g(x)\ . \tag{2.21}$$

But clearly this is the case for any real λ (to be precise, we could take $g(x)$ to be an indicator function for a finite interval including $x = \lambda$). So for X the spectrum is the entire real line.

This is intuitively compatible with our Postulate III, since the possible observable values of the position operator should be roughly the entire real line. However now there is apparently some tension with our desire to assign a basis of eigenstates to the set of points in the spectrum of an observable. Dirac suggested in his original treatise on the subject to forge ahead and formally extend his bra-ket formalism to include kets associated even to elements of a continuous spectrum. This is indeed the approach method that is standard in the physics community. His proposal can in retrospect be understood as being essentially an application of the spectral theorem for self-adjoint operators in its most sophisticated form. We will introduce the method now in an operational sense.

2.3.1 Generalised position eigenstates

To get our discussion off the ground, let's continue with the particle moving on the real line, so with Hilbert space $L^2(\mathbb{R})$.[20] The two fundamental observables in this setting are the position and momentum operators, and as we saw above, for the position operator X the spectrum is the entire real line.

Dirac instructs us to define an *generalised position eigenstate* $|\xi\rangle$ for this operator for each $\xi \in \mathbb{R}$,[21]

$$X\,|\xi\rangle = \xi\,|\xi\rangle \ . \tag{2.22}$$

Were we to use a wave function $\psi_\xi(x)$ to represent such a state, it would have to satisfy the unlikely-looking identity

$$x\psi_\xi(x) = \xi\psi_\xi(x) \ . \tag{2.23}$$

For this to hold, it must be that $\psi_\xi(x) = 0$ for $x \neq \xi$, and indeed if this were an element of $L^2(\mathbb{R})$ that would mean it was the zero function, so certainly this can't correspond to a non-zero element of the Hilbert space.

Nevertheless, we formally introduce such an object. Since this generalised state is meant to represent a situation where the particle is *definitely at* $x = \xi$, it is reasonable to demand

$$\langle\xi|\psi\rangle = \psi(\xi) \ , \qquad \langle\psi|\xi\rangle = \overline{\psi(\xi)} \ . \tag{2.24}$$

This is actually an important idea: *the value of the wave function at a point $x = \xi$ is the overlap of the state in question with the generalised position eigenstate $|\xi\rangle$*. Expressing this in terms of wave functions, we have

$$\int_{-\infty}^{\infty} \mathrm{d}x\,\overline{\psi_\xi(x)}\psi(x) = \psi(\xi) \ , \tag{2.25}$$

We recognise this to be precisely the *sifting property* of the (confusingly named) *Dirac δ-function*. Rather than a function, this is a *distribution*, meaning it is a linear functional on functions. You have met the Dirac δ-function previously in **M4 Multivariable Calculus**, and maybe also in **ASO Integral Transforms**. Indeed, we will identify

$$|\xi\rangle \ \longleftrightarrow \ \psi_\xi(x) = \delta(x-\xi) \ . \tag{2.26}$$

Note that while these generalised position eigenstates are not normalisable in the usual sense of $L^2(\mathbb{R})$, they do obey a *continuum normalisation condition*,[22]

$$\langle\xi|\xi'\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\delta(x-\xi)\delta(x-\xi') = \delta(\xi-\xi') \ . \tag{2.27}$$

This is a fairly natural generalisation of the usual normalisation condition where we have a Kronecker δ, but with the Dirac δ instead.

Happily, it turns out that we can for the most part use these generalised position eigenstates in the same ways we would use ordinary basis states as discussed previously, with various sums converted into integrals as appropriate. Justification

---

[20] A similar discussion here could take place for the particle moving on an interval $[0,1] \subset \mathbb{R}$, with Hilbert space $L^2([0,1])$. The free particle on the entire real line is even a bit more subtle.

[21] Here we begin to adopt a fairly standard notational choice: in the context of discussing a particular observable (in this case X), we denote states whose eigenvalue is some number (in this case $\xi \in \mathbb{R}$) by a *ket* whose label is *that same eigenvalue* (in this case $|\xi\rangle$). There is some danger of getting confused if not sufficiently diligent with this notational system, so be careful!

[22] Such (generalised) states are sometimes referred to as being δ-function normalisable states.

for this rests upon some deep pieces of functional analysis that we are sweeping under the rug,[23] but as we mentioned above, the quantum mechanical formalism (due to Dirac) actually predated the rigorous justification. In particular, we have a resolution of the identity in terms of these position eigenstates,

$$1_{L^2(\mathbb{R})} = \int_{-\infty}^{\infty} \mathrm{d}\xi \; |\xi\rangle \langle\xi| \; . \tag{2.28}$$

Acting on genuine states (corresponding to authentic, normalisable wave functions), we have

$$\begin{aligned} 1_{L^2(\mathbb{R})} |\psi\rangle &= \int_{-\infty}^{\infty} \mathrm{d}\xi \; |\xi\rangle \langle\xi|\psi\rangle \; , \\ &= \int_{-\infty}^{\infty} \mathrm{d}\xi \; \psi(\xi) \, |\xi\rangle \; , \end{aligned} \tag{2.29}$$

The final expression gives the continuum analogue of the decomposition of a general state in an orthonormal basis.

Generalising this resolution of the identity, if we integrate the outer product $|\xi\rangle\langle\xi|$ over any measurable subset $E \subset \mathbb{R}$, we obtain the self-adjoint projection operator corresponding to multiplication by the indicator function 1_E discussed in Chapter 1.3,[24]

$$\Pi_E = \int_E \mathrm{d}\xi \; |\xi\rangle \langle\xi| \; , \qquad \langle x| \, \Pi_E \, |\psi\rangle = 1_E(x)\psi(x) \; . \tag{2.30}$$

Indeed, we note that for a finite measure subset E this gives an actual projection operator on $\mathcal{H}$, while the outer product itself is not well-defined as a map on the Hilbert space. Since these projections are supposed to arise when considering measurements, this state of affairs is often understood as corresponding to the physical impossibility of measuring position with infinite precision; one could only ever check that a particle is within some error bar of a particular position.

2.3.2 Generalised momentum eigenstates

There is a similar story with the momentum operator $P = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}$. We introduce (generalised) momentum eigenstates $|p\rangle$,

$$P\,|p\rangle = p\,|p\rangle \; , \qquad p \in \mathbb{R} \, , \tag{2.31}$$

and if we denote the actual wave function associated to this state as $\psi_p(x)$, then we can easily solve the corresponding differential equation, at least formally,

$$-i\hbar\,\psi_p'(x) = p\psi_p(x) \qquad \Longrightarrow \qquad \psi_p(x) = \mathcal{N} e^{\frac{ipx}{\hbar}} \; , \tag{2.32}$$

where $\mathcal{N}$ is some normalisation factor. The problem is clear and feels familiar: these wave functions are not square-normalisable at all (on the entire real line), so this is not giving us an element of $L^2(\mathbb{R})$. Rather this is a kind of generalised eigenstate, which we can again interpret as a distribution.

Using our previous understanding of the relationship between wave functions and generalised position eigenstates, we deduce the overlap equation

$$\langle x|p\rangle = \psi_p(x) = \mathcal{N} e^{\frac{ipx}{\hbar}} \; . \tag{2.33}$$

---

[23] There are several realisations of these generalised eigenstates within a more rigorous framework. In one version of the spectral theorem for self-adjoint operators on infinite-dimensional Hilbert spaces, one constructs the Hilbert space of interest as a *direct integral* of smaller Hilbert spaces, and these generalised states can be understood as elements of the (Hilbert-space) integrand of that direct integral. Alternatively, Hilbert spaces arising in quantum mechanics can be equipped with additional structure known as a *Gel'fand triple*. In this case the generalised states are elements of a larger space of distributions that form a part of that structure. You don't need to know any of this for the present course, but it is a beautiful subject!

[24] In one rigorous treatment of these constructions, it is this assignment of a self-adjoint projection operator to (measurable) subsets of $\mathbb{R}$ that is rigorously defined and guaranteed to exist by the spectral theorem; such an assignment is called a *projection valued measure*.

We can then derive the continuum normalisation condition for the momentum eigenstates,

$$\langle p|p'\rangle = |\mathcal{N}|^2 \int_{-\infty}^{\infty} \mathrm{d}x\, e^{-\frac{ipx}{\hbar}} e^{\frac{ip'x}{\hbar}} = 2\pi\hbar\,|\mathcal{N}|^2 \int_{-\infty}^{\infty} \mathrm{d}s\, e^{2\pi i(p-p')s} = 2\pi\hbar\,|\mathcal{N}|^2\delta(p-p')\,, \tag{2.34}$$

where in the last equation we have recognised the integral representation for the delta function. It is then natural to adopt the normalisation conventions $\mathcal{N} = (2\pi\hbar)^{-1/2}$ giving canonical continuum normalisation to the generalised momentum eigenstates. We have an analogous resolution of the identity in terms of momentum states,

$$1_{\mathcal{H}} = \int_{-\infty}^{\infty} \mathrm{d}p\ |p\rangle\langle p|\ . \tag{2.35}$$

This formalism of position and momentum (generalised) bases for $L^2(\mathbb{R})$ gives us a nice new perspective on the quantum mechanics of a particle. To a given state vector $|\psi\rangle$, we can associated either its expression in *position space*,

$$\psi(x) = \langle x|\psi\rangle\ , \qquad |\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\psi(x)\,|x\rangle\ , \tag{2.36}$$

or its expression in *momentum space*,

$$\widehat{\psi}(p) = \langle p|\psi\rangle\ , \qquad |\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\,\widehat{\psi}(p)\,|p\rangle\ . \tag{2.37}$$

So there are actually (at least) two wave functions associated to the state ψ of a particle on the real line. These, it turns out, are related by the Fourier transform,

$$\psi(x) = \langle x|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\ \langle x|p\rangle\langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} \mathrm{d}p\, e^{\frac{ipx}{\hbar}}\,\widehat{\psi}(p)\ . \tag{2.38}$$

$$\widehat{\psi}(p) = \langle p|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\ \langle p|x\rangle\langle x|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} \mathrm{d}x\, e^{-\frac{ipx}{\hbar}}\,\psi(x)\ . \tag{2.39}$$

Indeed, the Fourier transform is a unitary map from $L^2(\mathbb{R})$ to itself (this is the Plancherel theorem), so the change from position to momentum representation is just *a change of basis* for our Hilbert space.

Remark 2.3.2. This formalism for generalised position eigenstates generalises immediately to the case of a particle moving in, say, $d = 2$ or $d = 3$ dimensions. There for $\mathbf{x} = (x_1, \ldots, x_d)$ we have the generalised eigenstates

$$X_i\,|\mathbf{x}\rangle = x_i\,|\mathbf{x}\rangle\ , \tag{2.40}$$

obeying the continuum normalisation condition,

$$\langle\mathbf{x}|\mathbf{x}'\rangle = \delta^d(\mathbf{x}-\mathbf{x}')\,, \tag{2.41}$$

and the corresponding resolution of the identity,

$$1_{L^2(\mathbb{R}^d)} = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \mathrm{d}x_1\cdots\mathrm{d}x_d\ |\mathbf{x}\rangle\langle\mathbf{x}|\ . \tag{2.42}$$

Similarly, we have generalised momentum eigenstates corresponding to non-normalisable plane-wave wave functions,

$$|\mathbf{p}\rangle \quad \longrightarrow \quad \psi_{\mathbf{p}}(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{d/2}} e^{\frac{i\mathbf{p}\cdot\mathbf{x}}{\hbar}} , \tag{2.43}$$

obeying the same continuum normalisation condition and admitting the same type of resolution of the identity. d-dimensional wave functions in position space and momentum space are related now by the d-dimensional Fourier transform.

2.4 Application: free particle propagator

A nice application of the machinery we have developed here is in defining an important object in studying quantum mechanical dynamics: the *propagator*. Intuitively, this is the quantity that tells you the quantum mechanical *amplitude* (square root of probability density) for a particle that starts at a given position to be detected at some other position at some definite time in the future. In terms of generalised position eigenstates, this is the quantity

$$U(x_1, t_1; x_0, t_0) := \langle x_1 | U(t_1; t_0) | x_0 \rangle \ , \tag{2.44}$$

where $U(t_1; t_0)$ is the unitary time evolution operator introduced previously. If one has a reasonably good understanding of the propagator, then the time evolution of general quantum states can be described using the following double integral

$$\begin{aligned} \langle \psi_1 | U(t_1, t_0) | \psi_0 \rangle &= \langle \psi_1 | \left(\int_{-\infty}^{\infty} \mathrm{d}x_1 \ |x_1\rangle \langle x_1| \right) U(t_1; t_0) \left(\int_{-\infty}^{\infty} \mathrm{d}x_0 \ |x_0\rangle \langle x_0| \right) |\psi_0\rangle \ , \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}x_1 \, \mathrm{d}x_0 \ \overline{\psi_1(x_1)} U(x_1, t_1; x_0, t_0) \psi_0(x_0) \ . \end{aligned} \tag{2.45}$$

In general, the propagator is not so easy to compute. Here we will do it for the case of the free particle in one dimension. The Hamiltonian is $H = P^2/2m$ and the (generalised) energy eigenstates are precisely the (generalised) momentum eigenstates:

$$H |p\rangle = E_p |p\rangle = \frac{p^2}{2m} |p\rangle \ . \tag{2.46}$$

As we know well, time evolution for these states then proceeds via phase multiplication,

$$U(t_1, t_0) |p\rangle = \exp\left(\frac{-iE_p(t_1 - t_0)}{\hbar} \right) |p\rangle = \exp\left(\frac{-ip^2(t_1 - t_0)}{2m\hbar} \right) |p\rangle \ . \tag{2.47}$$

This means that the *momentum-space propagator* is very simple for the free particle,

$$\langle p_1 | \, U(t_1; t_0) \, |p_0\rangle =: \widehat{U}(p_1, t_1; p_0, t_0) = \delta(p_1 - p_0) \exp\left(-\frac{ip_0^2(t_1 - t_0)}{2m\hbar} \right) \ . \tag{2.48}$$

The position-space propagator is then obtained by a double Fourier transform,[25]

$$U(x_1, t_1; x_0, t_0) = \quad \langle x_1 | \left(\int_{-\infty}^{\infty} \mathrm{d}p_0 \, |p_0\rangle \, \langle p_0| \right) U(t_1; t_0) \left(\int_{-\infty}^{\infty} \mathrm{d}p_1 \ |p_1\rangle \, \langle p_1| \right) |x_0\rangle \ , \tag{2.49}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}p_0 \, \mathrm{d}p_1 \ \langle x_1 | p_1 \rangle \ \widehat{U}(p_1, t_1; p_0, t_0) \ \langle p_0 | x_0 \rangle \ , \tag{2.50}$$

---

[25] The attentive reader will notice that upon setting $t \to -i\tau$, the final result for the propagator becomes the Green's function for the one-dimensional heat equation, with the thermal conductivity given by $\hbar/2m$. Indeed, the same "imaginary time" replacement applied to the time-dependent Schrödinger equation for this system yields the heat equation with said thermal conductivity.

$$= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}p_0\,\mathrm{d}p_1\,\delta(p_1-p_0)\exp\left(\frac{ip_1x_1-ip_0x_0}{\hbar}-\frac{ip_0^2(t_1-t_0)}{2m\hbar}\right), \tag{2.51}$$

$$= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \mathrm{d}p\,\exp\left(\frac{ip(x_1-x_0)}{\hbar}-\frac{ip^2(t_1-t_0)}{2m\hbar}\right), \tag{2.52}$$

$$= \left(\frac{m}{2\pi i\hbar(t_1-t_0)}\right)^{\frac{1}{2}}\exp\left(-\frac{m(x_1-x_0)^2}{2i\hbar(t_1-t_0)}\right). \tag{2.53}$$

The last integral is somewhat subtle, but can be computed using results for *Fresnel integrals*.[26]

It is interesting to observe that instantly when $t_1 > t_0$, the propagator is nonzero for arbitrarily large $x_1 - x_0$. This reflects the infinite uncertainty in momentum that is associated with the completely localised position eigenstate at time t_0. However, the phase in the exponential is also very large for large $x_1 - x_0$ and small $t_1 - t_0$, so when we average over positions (as we should if we start with a normalisable wave function) then there will be cancellations and the wave function will remain somewhat localised near its original support.

---

[26] The general formula is

$$\int_{-\infty}^{\infty} \mathrm{d}x\,\exp\left(\frac{i}{2}ax^2+ibx\right)=\left(\frac{2\pi i}{a}\right)^{\frac{1}{2}}\exp\left(-\frac{ib^2}{2a}\right).$$

Chapter 3

Composite Systems, Tensor Products, Entanglement

It is important to be able to build up the description of a quantum system from those of more elementary subsystems with a smaller number of degrees of freedom. In this section we introduce the basic mathematical machinery for doing this (the Hilbert space tensor product), and look at simple aspects of the phenomenon of quantum entanglement, which arises naturally as a consequence.

3.1 Hilbert space tensor product

Suppose we encounter two quantum systems that (at least in some idealisation) do not interact with one another, and that are to be taken together in a single description. (You might imagine two atoms kept far enough apart so as to be non-interacting.) The two systems, taken on their own, will have their states encoded by Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, while as a composite system we should assign a single Hilbert space $\mathcal{H}_3$. How should this Hilbert space $\mathcal{H}_3$ be characterised? The following construction arises naturally from physical considerations:

- For state vectors $|\psi_1\rangle \in \mathcal{H}_1$ and $|\psi_2\rangle \in \mathcal{H}_2$, there should exist a definite state vector that we will denote by $|\psi_1 \otimes \psi_2\rangle \in \mathcal{H}_3$. Such a vector is referred to as a *pure tensor* or a *simple tensor*, or alternatively, as *decomposable*.[27] The set of objects defined in this manner is just $\mathcal{H}_1 \times \mathcal{H}_2$.
- By the linearity of quantum theory, we should be able to take linear superpositions of these pure tensors.

At this stage we have effectively reproduced the following definition:

Definition 3.1.1 (Free Vector Space)**.** The *free vector space* on the set $\mathcal{H}_1 \times \mathcal{H}_2$ is the vector space of all finite linear combinations of elements of $\mathcal{H}_1 \times \mathcal{H}_2$.

The free vector space overcounts in some obvious ways, and we introduce a number of identifications.

- Since state vectors only encode physical states up to overall scalar multiplication, the consequence of rescaling either tensor factor should be no different from rescaling the vector as a whole:

$$|\lambda\psi_1 \otimes \psi_2\rangle \sim \lambda\,|\psi_1 \otimes \psi_2\rangle \sim |\psi_1 \otimes \lambda\psi_2\rangle \ \text{ for } \lambda \in \mathbb{C}\,.$$

- If system two is definitely in state $|\psi_2\rangle$, then when system one is in a superposition of two states, the total system is in the superposition of the corresponding two decomposable states where the second system remains in $|\psi_2\rangle$.

$$|(\psi_1 + \varphi_1) \otimes \psi_2\rangle \sim |\psi_1 \otimes \psi_2\rangle + |\varphi_1 \otimes \psi_2\rangle \ .$$

- The same argument as above should hold with the two systems switched.

$$|\psi_1 \otimes (\psi_2 + \varphi_2)\rangle \sim |\psi_1 \otimes \psi_2\rangle + |\psi_1 \otimes \varphi_2\rangle \ .$$

Definition 3.1.2 (Tensor Product)**.** The *vector space tensor product* is then defined as the quotient of the free vector space above by the equivalence relations given above,

$$\mathcal{H}_1 \otimes \mathcal{H}_2 := F(\mathcal{H}_1, \mathcal{H}_2)/\sim \ . \tag{3.1}$$

---

[27] When not using *bra-ket* notation, it is common to write this vector as $\psi_2 \otimes \psi_2$. We may also sometime use $|\psi_1, \psi_2\rangle$ or $|\psi_1\rangle \otimes |\psi_2\rangle$ interchangeably. Hopefully the meaning of the notation will always be clear in context.

An inner product on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is inherited from those on $\mathcal{H}_1$ and $\mathcal{H}_2$ by defining for pure tensors

$$\langle \psi_1 \otimes \psi_2 | \varphi_1 \otimes \varphi_2 \rangle = \langle \psi_1 | \varphi_1 \rangle \langle \psi_2 | \varphi_2 \rangle \ , \tag{3.2}$$

and extending this to general elements by sesquilinearity. Note that this definition is compatible with the equivalence relations given above.

Definition 3.1.3 (Hilbert Space Tensor Product)**.** The Hilbert space tensor product is obtained by taking the *completion* of this vector space tensor product with respect to the norm induced by the inner product.

As usual, this completion is a technical detail that is relevant in the infinite dimensional case. It essentially means that we allow the possibility of taking infinite linear combinations of pure tensors whose norm is still finite. Finer aspects of this construction won't be important or examinable in this course, though we will see examples.

Remark 3.1.4 (Alternate construction of tensor product). There is an equivalent, in a sense much simpler, definition of the Hilbert space tensor product that is often used. Let $\{|\alpha_i\rangle\}$ and $\{|\beta_j\rangle\}$ denote bases for $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Then $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be identified with the Hilbert space with given basis $\{|\alpha_i \otimes \beta_j\rangle\}$ (again, in the infinite dimensional case one requires completeness, which allows infinite linear combinations of these with finite norm). To a purist, the first definition has the advantage being explicitly basis-independent. For practical purposes, this latter definition is often the most useful.

If a system is described by a tensor product Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$, then operators and observables that are defined to act separately on $\mathcal{H}_1$ and $\mathcal{H}_2$ naturally extend to the tensor product. If $A_1 : \mathcal{H}_1 \to \mathcal{H}_1$ and $A_2 : \mathcal{H}_2 \to \mathcal{H}_2$, then we can define

$$\begin{aligned} A_1 \otimes A_2 : \mathcal{H}_1 \otimes \mathcal{H}_2 &\longrightarrow \mathcal{H}_1 \otimes \mathcal{H}_2 \\ |\psi_1 \otimes \psi_2\rangle &\longmapsto |A_1\psi_1 \otimes A_2\psi_2\rangle \ . \end{aligned} \tag{3.3}$$

In particular, when either A_1 or A_2 is the identity operator, then this gives operators that act on the tensor product only through the second or first tensor factor, respectively. Such operators naturally commute,

$$(A_1 \otimes 1_{\mathcal{H}_2})(1_{\mathcal{H}_1} \otimes A_2) = (A_1 \otimes A_2) = (1_{\mathcal{H}_1} \otimes A_2)(A_1 \otimes 1_{\mathcal{H}_2}) \ , \tag{3.4}$$

which is in agreement with the physical criterion that making observations on one system should not impact another, in principle disjoint, system.

We can similarly form the n-fold tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n$ with basis $\alpha_{i_1} \otimes \beta_{i_2} \otimes \ldots \otimes \gamma_{i_n}$ with i_j indexing a basis of $\mathcal{H}_j$. This is the Hilbert space for the composite of the n quantum mechanical systems described by $\mathcal{H}_i$, $i = 1, \ldots, n$. When the constituent Hilbert spaces are all identical to $\mathcal{H}$ we simply write $\otimes^n \mathcal{H}$ or $\mathcal{H}^{\otimes n}$.

A first important behaviour of Hilbert spaces under tensor product is that their dimensions (when finite) combine multiplicatively,

$$\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim \mathcal{H}_1 \times \dim \mathcal{H}_2 \ . \tag{3.5}$$

This follows immediately from the second construction of the tensor product given above, where the number of basis elements clearly obeys this relation. It is worth pausing to compare this situation with what one encounters classically. If a two classical systems have configuration spaces of dimensions d_1 an d_2, say, then taken together their joint configuration space will be of dimension $d_1 + d_2$. In this sense, quantum mechanical state spaces get very big very fast compared to their classical analogues. Indeed, this is one of the properties that underlies the power of quantum computation.

There is a subspace of the tensor product Hilbert space that behaves a bit more classically: this is the subspace of pure tensors. Note that this is not a linear subspace of $\mathcal{H}_1 \otimes \mathcal{H}_2$, since the property of being a pure tensor is not preserved under addition. The dimensionality of the subspace of pure tensors does behave additively,

$$\dim(\mathcal{H}_1 \otimes \mathcal{H}_2)_{\text{decomposable}} = \dim \mathcal{H}_1 + \dim \mathcal{H}_2 - 1 \ , \tag{3.6}$$

where the correction by one comes from the equivalence of rescaling the two tensor factors in a pure tensor.

Remark 3.1.5. Another way to see this result is to consider the relevant subspaces in projectivised Hilbert space. Here we have that subspace of decomposable states is of the form

$$\mathbb{P}(\mathcal{H}_1) \times \mathbb{P}(\mathcal{H}_2) \subset \mathbb{P}(\mathcal{H}_1 \otimes \mathcal{H}_2)\,. \tag{3.7}$$

The dimensionality of the left hand side is $\dim \mathcal{H}_1 + \dim \mathcal{H}_2 - 2$, and deprojectivising to recover the subspace of the Hilbert space adds one dimension. The embedding describing how the left hand side of (3.7) sits inside the right hand side is known as the *Segre embedding*.

We see that in composite systems, the pure tensors are highly non-generic. The generic state, on the other hand, displays what is called *quantum entanglement*.

Definition 3.1.6 (Entangled State). In a composite quantum system, a state which is not described by a pure tensor is known as an entangled state.

Below we explore the notion of entanglement a bit in a simple example: the composition of two qubits.

3.2 Tensor product of qubits; entanglement

Let's consider the tensor product in the simplest case of combining several qubits (see Chapter 1.2). We recall that the qubit has Hilbert space $\mathcal{H} \cong \mathbb{C}^2$; let us now (adopting Dirac notation) fixed an orthonormal basis $\{|0\rangle, |1\rangle\}$ for the qubit such that $\sigma_3 |1\rangle = 1$ and $\sigma_3 |0\rangle = -1$.[28] We can then take as a basis for the tensor product $\mathbb{C}^2 \otimes \mathbb{C}^2 \cong \mathbb{C}^4$ the following pure tensors

$$|0 \otimes 0\rangle\ , \quad |0 \otimes 1\rangle\ , \quad |1 \otimes 0\rangle\ , \quad |1 \otimes 1\rangle\ . \tag{3.8}$$

Within this vector space, the most general state takes the form

$$a\,|0 \otimes 0\rangle + b\,|0 \otimes 1\rangle + c\,|1 \otimes 0\rangle + d\,|1 \otimes 1\rangle\ . \tag{3.9}$$

while the most general pure tensor takes the form

$$(\alpha\,|0\rangle + \beta\,|1\rangle) \otimes (\gamma\,|0\rangle + \delta\,|1\rangle) = \alpha\gamma\,|0 \otimes 0\rangle + \alpha\delta\,|0 \otimes 1\rangle + \beta\gamma\,|1 \otimes 0\rangle + \beta\delta\,|1 \otimes 1\rangle \tag{3.10}$$

One can check that a state of the form (3.9) can be written as in (3.10) if and only if $ad - bc = 0$, so indeed the set of pure states is a nonlinear subspace of $\mathcal{H}$ with dimension $2 + 2 - 1 = 3$.

If we combine more qubits the dimension of the Hilbert space grows exponentially. In particular,

$$\otimes^n \mathbb{C}^2 \cong \mathbb{C}^{2^n}\ , \tag{3.11}$$

while the space of pure tensors is dramatically smaller (namely $2n - 1$). Indeed, entangled states are by a wide margin the generic ones in large composite quantum systems.

The two qubit system lets us study some of the strange properties of quantum entangled systems. Suppose we have two qubit systems that are prepared (somehow) in the initial state[29]

$$|\mathrm{EPR}\rangle = \frac{|1 \otimes 0\rangle - |0 \otimes 1\rangle}{\sqrt{2}}\ . \tag{3.12}$$

---

[28] These basis vectors are often denoted instead by $|\uparrow\rangle$ and $|\downarrow\rangle$ respectively due to their interpretation in terms of spins, and sometimes also $|+\rangle$ and $|-\rangle$. We may use either or both of these when we revisit this system in later chapters.

[29] EPR here stands for Einstein–Podolsky–Rosen, the authors of a famous paper pointing out seemingly paradoxical properties of entangled quantum systems. This kind of a state is also sometimes called a *Bell pair*.

Suppose now that Alice carries the first qubit with her to a faraway star system, while Bob remains on Earth with the second qubit. If Alice takes a measurement corresponding to the observable σ_3 on her qubit (so corresponding to the observable $\sigma_3 \otimes \sigma_0$ on the tensor product Hilbert space), there is a 50% probability that she will find the value $+1$ and a 50% probability that she will find the value -1. In either case, she should find that the quantum state collapses according to the wave function collapse postulate,

$$\begin{aligned} &\text{Alice measures } \sigma_3 \text{ finds } +1 \quad \Longrightarrow \quad |\psi\rangle = |1 \otimes 0\rangle \ , \\ &\text{Alice measures } \sigma_3 \text{ finds } -1 \quad \Longrightarrow \quad |\psi\rangle = |0 \otimes 1\rangle \ . \end{aligned} \tag{3.13}$$

In each of the collapsed states, the results of a σ_3 measurement by Bob of his qubit (corresponding to the observable $\sigma_0 \otimes \sigma_3$ on the combined system) should return a definite answer. Which answer is returned is dictated by the results of Alice's measurement. One might phrase this in a paradoxical-sounding way, as saying that when Alice makes her measurement, it instantaneously impacts the outcomes of Bob's experiment.

A still more surprising version of this situation occurs if we consider the possibility that Alice might either measure σ_3 or, say, σ_1, while Bob will definitely measure σ_3. In the former case, as we said above, the result of Bob's experiment is determined completely once Alice's measurement has been performed. However, if Alice performs a σ_1 measurement, then the resulting state after acting with the appropriate projection operator is

$$\begin{aligned} &\text{Alice measures } \sigma_1 \text{ finds } +1 \quad \longrightarrow \quad |\psi\rangle \sim |1 \otimes 1\rangle + |0 \otimes 1\rangle - |0 \otimes 0\rangle - |1 \otimes 0\rangle \ , \\ &\text{Alice measures } \sigma_1 \text{ finds } -1 \quad \longrightarrow \quad |\psi\rangle \sim |1 \otimes 1\rangle + |1 \otimes 0\rangle - |0 \otimes 0\rangle - |0 \otimes 1\rangle \ . \end{aligned} \tag{3.14}$$

In this case, depending on Alice's choice of what observable to measure, the probability distribution of outcomes for Bob's measurement changes completely. This sounds odd, especially in view of Einstein's theory of relativity, which says that there should be no communication faster than the speed of light. Upon additional scrutiny, however, the situation is perhaps not quite so paradoxical; though the result of Alice's measurement (and even her choice of what to measure) has an implication for what Bob might measure, Bob has no way of knowing what result Alice done or found. What we really get from the entangled state is an interesting set of *correlations* between the results of various experiments Alice and Bob might peform.[30]

3.3 Multi-particle systems of distinguishable particles

Another incarnation of the tensor product arises when we consider systems of several elementary particles. If our particles move in d dimensional space, then the Hilbert space for the i'th particle will be identified as $\mathcal{H}_i \cong L^2(\mathbb{R}^d)$, and for n particles we are supposed to be interested in the Hilbert space

$$\mathcal{H} \cong L^2(\mathbb{R}^d)_1 \otimes L^2(\mathbb{R}^d)_2 \otimes \cdots \otimes L^2(\mathbb{R}^d)_n \ . \tag{3.15}$$

The result of the Hilbert space tensor product turns out to just be the space of square-normalisable wave functions of the n particle positions $\psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$, *i.e.*,

$$\mathcal{H} \cong L^2(\mathbb{R}^{d\times n}) \ . \tag{3.16}$$

At a technical level, this is a case where the final step of *completing* the Hilbert space is relevant. We identify a pure tensor of single-particle wave functions with a separable n-particle wave function,

$$\psi_1(\mathbf{x}_1) \otimes \psi_1(\mathbf{x}_1) \otimes \cdots \otimes \psi_n(\mathbf{x}_n) \quad \longleftrightarrow \quad \psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \psi_1(\mathbf{x}_1)\psi_1(\mathbf{x}_1) \cdots \psi_n(\mathbf{x}_n) \ . \tag{3.17}$$

A general n-particle wave function certainly can't be written as a finite linear combination of separable wave functions of the above form. However, given a basis $\psi_i(\mathbf{x})$, $i = 1, 2, \ldots, \infty$ for $L^2(\mathbb{R}^d)$, pure tensors formed from these basis elements do form an orthonormal basis for $L^2(\mathbb{R}^{d\times n})$.

---

[30] There is a lot more to say here and we won't pursue it in this course. A further refinement of this hypothetical, due to John Stewart Bell, led to the famed *Bell's inequality*, which highlights the degree to which quantum physics diverges from what is possible in a classical world. The subject is worth investigating for one's own edification.

Remark 3.3.1. The technical subtlety associated with completion of the Hilbert space is, at least formally, evaded when we choose to work with generalised position eigenstates. In this case, we introduce basis elements

$$|\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle = |\mathbf{x}_1\rangle \otimes \cdots \otimes |\mathbf{x}_n\rangle \ , \tag{3.18}$$

which obey

$$\langle \mathbf{x}_1, \ldots, \mathbf{x}_n | \mathbf{x}_1', \ldots, \mathbf{x}_n' \rangle = \delta^d(\mathbf{x}_1 - \mathbf{x}_1') \cdots \delta^d(\mathbf{x}_n - \mathbf{x}_n') = \delta^{n\times d}((\mathbf{x}_1; \ldots; \mathbf{x}_n) - (\mathbf{x}_1'; \ldots; \mathbf{x}_n')) \ , \tag{3.19}$$

where in the last expression we are using the $n \times d$-dimensional Dirac delta function. Then the most general state takes the form

$$|\psi\rangle = \int_{\mathbb{R}^{nd}} \mathrm{d}^d\mathbf{x}_1 \cdots \mathrm{d}^d\mathbf{x}_n \psi(\mathbf{x}_1, \ldots, \mathbf{x}_n) \, |\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle \ , \tag{3.20}$$

which is just an n-particle wave function in the usual sense as an element of $L^2(\mathbb{R}^{n\times d})$.

Chapter 4

Identical Particles and Statistics

A curious consequence of the quantum mechanical picture of the world is that elementary particles of the same type (electrons, quarks, *etc.*) naturally come to be thought of as being *fundamentally indistinguishable*. To motivate this consider the following thought experiment (see Figure 1).[31]

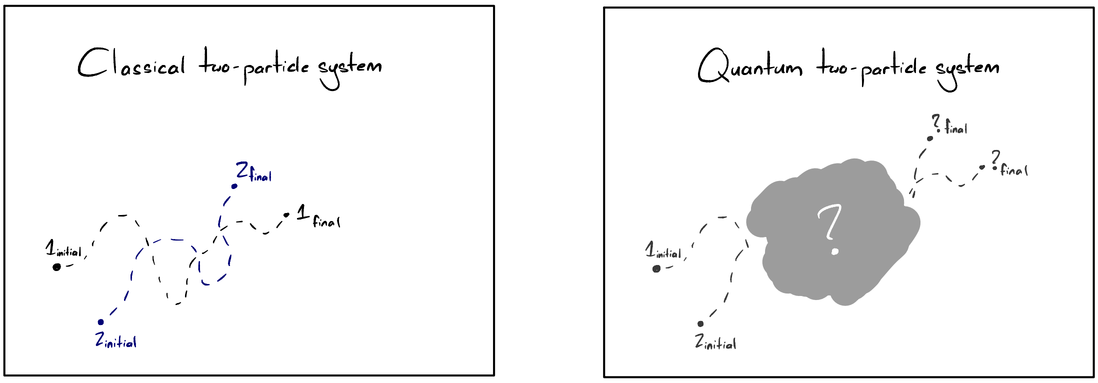


Figure 1. Distinguishable versus indistinguishable elementary particles in classical and quantum physics.

Suppose that at time $t = 0$ you have a pair of electrons whose positions are known well enough to distinguish them (*i.e.*, one is definitely in one half of the room and the other is definitely in the other half). After a period of time elapses, the wave function of the system will have evolved so that both electrons could be anywhere in the room with some probability. At this point, we may make measurements to determine the positions of two electrons, but we will have no way to distinguish the two different electrons. In contrast to the classical case, we can't keep track of "electron one" by following it along its trajectory during the time interval—it had no definite trajectory when it was not being observed!

4.1 Indistinguishable particles and wave functions

We will consider the consequences of indistinguishability on multi-particle wave functions, beginning with the case of two particles. A two particle wave function is a (square-normalisable) function of two positions $\psi(\mathbf{x}_1, \mathbf{x}_2)$. If we now demand that the wave function represent indistinguishable particles, then it should assign the same probability (density) to find "particle one" at $\mathbf{x}_1$ and "particle two" at $\mathbf{x}_2$ as it does to finding "particle one" at $\mathbf{x}_2$ and "particle two" at $\mathbf{x}_1$.[32] In other words, the wave function should obey

$$|\psi(\mathbf{x}_2, \mathbf{x}_1)|^2 = |\psi(\mathbf{x}_1, \mathbf{x}_2)|^2 \quad \Longrightarrow \quad \psi(\mathbf{x}_2, \mathbf{x}_1) = \lambda\,\psi(\mathbf{x}_1, \mathbf{x}_2)\ , \tag{4.1}$$

where $\lambda = e^{i\phi}$ is a *phase*. Iterating this relation we see that

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \lambda^2\psi(\mathbf{x}_1, \mathbf{x}_2)\ , \tag{4.2}$$

so there are just two possibilities: $\lambda = \pm 1$.[33] In the case $\lambda = 1$ we are restricting ourselves to symmetric functions of the two particles' positions, while for $\lambda = -1$ we have anti-symmetric functions.

---

[31] It also turns out that the indistinguishability of elementary particles is built directly into *quantum field theory*, which is the framework that synthesises quantum theory with the special theory of relativity.

[32] Scare quotes because, of course, there is no unambiguous notion of particle one and particle two; this is just referring to the order of the arguments in the wave function.

[33] There is an oft-mentioned caveat here, which is that in two-dimensions there is a possibility for a more general phase λ, with the corresponding particles referred to as *anyons*. To see the possibility of this more general phase, it is necessary to be a bit more flexible about the description of the Hilbert space for several particles to allow multi-valued functions of positions. We will not pursue this here.

4.1.1 Permutations and Statistics

In the case of many particles, the exchange of the two particle positions generalises to an action of the symmetric group S_n on the space of n-particle wavefunctions,

$$\begin{aligned} \big(S_n, L^2(\mathbb{R}^{n\times d})\big) &\longrightarrow L^2(\mathbb{R}^{n\times d}) \\ \big(\pi, \psi(\mathbf{x}_1,\dots,\mathbf{x}_n)\big) &\longmapsto (\pi\circ\psi)(\mathbf{x}_1,\dots,\mathbf{x}_n) = \psi(\mathbf{x}_{\pi(1)},\dots,\mathbf{x}_{\pi(n)})\ . \end{aligned} \tag{4.3}$$

By the same argument we had for the case $n = 2$, indistinguishability implies that this action obeys

$$\psi(\mathbf{x}_{\pi(1)},\dots,\mathbf{x}_{\pi(n)}) = \lambda(\pi)\psi(\mathbf{x}_1,\dots,\mathbf{x}_n)\ , \tag{4.4}$$

with $\lambda(\pi)$ a (now π-dependent) phase: $\lambda(\pi) = e^{i\phi(\pi)}$. Composing the action of two permutations $\pi, \sigma \in S_n$ we find

$$\begin{aligned} \psi(\mathbf{x}_{(\pi\circ\sigma)(1)},\dots,\mathbf{x}_{(\pi\circ\sigma)(n)}) = \lambda(\pi)\psi(\mathbf{x}_{\sigma(1)},\dots,\mathbf{x}_{\sigma(n)}) &= \lambda(\pi)\lambda(\sigma)\psi(\mathbf{x}_1,\dots,\mathbf{x}_n)\ , \\ &= \lambda(\pi\circ\sigma)\psi(\mathbf{x}_1,\dots,\mathbf{x}_n)\ . \end{aligned} \tag{4.5}$$

So we have the rule $\lambda(\pi\circ\sigma) = \lambda(\pi)\lambda(\sigma)$.

Definition 4.1.1. A *multiplicative character* of a group G is a group homomorphism from G into the circle group $U(1)$ (or more generally into the ring of units $k^\times$ of a field k).

Thus, we have that the map $\lambda : S_n \to \mathbb{C}$ defines a multiplicative character for the permutation group S_n.

It turns out that there are only two inequivalent multiplicative characters for S_n. First observe that any two elements of S_n that are conjugate to each other are mapped to the same value by a multiplicative character,

$$\lambda(\pi\circ\sigma\circ\pi^{-1}) = \lambda(\pi)\lambda(\sigma)\lambda(\pi^{-1}) = \lambda(\pi)\lambda(\sigma)\lambda(\pi)^{-1} = \lambda(\sigma)\ , \tag{4.6}$$

where we have used that $\lambda(\pi^{-1}) = \lambda(\pi)^{-1}$, which follows from the character being a group homomorphism. Now recall that in S_n, a *transposition* is a permutation that just swaps two elements of $\{1,\dots,n\}$, say r and s, and is denoted $(r\,s)$. Such transpositions are all conjugate to one another:

$$(r\,s) = (1\,r)(2\,s)(1\,2)(2\,s)^{-1}(1\,r)^{-1}\ . \tag{4.7}$$

Thus we have that $\lambda((r\,s)) = \lambda((1\,2)) = \pm 1$, where our previous argument in the two-particle case implies the latter equality.

General permutations are generated by the composition of transpositions, and are unambiguously classified as either being *odd* or *even* according to whether they arise from an odd or even number of transpositions. Thus we have the following

Proposition 4.1.2. Let $\lambda : S_n \to \mathbb{C}$ be a multiplicative character for the symmetric group. Then either $\lambda(\pi) \equiv 1$ or $\lambda(\pi) = \varepsilon(\pi)$, where ε gives the signature of the permutation,

$$\varepsilon(\pi) := \begin{cases} 1 & \text{for } \pi \text{ even}\ , \\ -1 & \text{for } \pi \text{ odd}\ . \end{cases} \tag{4.8}$$

The two possibilities for wave functions of indistinguishable particles are then either *totally symmetric* wave functions ($\lambda \equiv 1$) or *totally antisymmetric* wave functions ($\lambda = \varepsilon$), generalising the two-particle case. For a given species of elementary (indistinguishable) particle, one of these two cases must apply. This leads to a binary classification of indistinguishable particles:

Definition 4.1.3. Indistinguishable particles satisfying (4.4) are called *bosons* if the corresponding group character is the trivial one; these particles are said to obey *Bose–Einstein statistics*. Particles satisfying (4.4) with the nontrivial character (ε) are called *fermions*; these particles are said to obey *Fermi–Dirac statistics*.

The known elementary fermions in nature are electrons, muons, τ-particles, and neutrinos, along with their anti-particles (collectively, *leptons*), as well as quarks. Also composite particles made up of an odd number of elementary fermions, such as protons and neutrons, are fermions. The known elementary bosons in nature are photons, gluons, W- and Z-bosons, gravitons, and the Higgs boson. Also composite particles made up of an even number of elementary fermions, such as *mesons*.

An important fact, which can be observed empirically in nature, is that the statistics of a particle is correlated with its *spin* (we will give a full treatment of spin in Chapter 5.2.5). In fact, this empirical fact is also a mathematical theorem that can be proven within the context of *relativistic quantum field theory*.

Theorem 4.1.4 (Spin-statistics theorem in three dimensions)**.** In a relativistic quantum theory in three spatial dimension, particles with integer spin must obey Bose–Einstein statistics. Particles with half-integer spin ($n + \frac{1}{2}$ for $n \in \mathbb{N}$) must obey Fermi–Dirac statistics.

An analogous theorem holds in any number of spatial dimensions greater than three, where one must be a bit more precise about the meaning of integer/half-integer spin (spin is no longer characterised by a single number in higher dimensions).

4.2 Bosonic and fermionic wave functions

It is useful to have practical tools for producing and manipulating wave functions for particles obeying appropriate statistics. To this end we can define projection operators onto the subspaces of completely symmetric and completely anti-symmetric (bosonic and fermionic, respectively) wavefunctions. In particular, for a general n-particle wave function ψ, define[34]

$$\Pi_\lambda \psi = \frac{1}{n!} \sum_{\pi \in S_n} \lambda(\pi^{-1}) \psi(\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(n)}) \,, \tag{4.9}$$

where as before, λ is the identity for Bose–Einstein and is ε for Fermi–Dirac. We can think of this as *averaging over the action of the permutation group*, with the average weighted by the relevant group character. We easily prove the following:

Proposition 4.2.1. For $\sigma \in S_n$ we have

$$(\Pi_\lambda \psi)(\mathbf{x}_{\sigma(1)}, \ldots, \mathbf{x}_{\sigma(n)}) = \lambda(\sigma)(\Pi_\lambda \psi)(\mathbf{x}_1, \ldots, \mathbf{x}_n) \,, \tag{4.10}$$

$\Pi_\lambda^2 = \Pi_\lambda$, and Π_λ is self-adjoint. Thus Π_λ is an orthogonal projection operator onto bosonic/fermionic wave functions.

Proof. For the first result, we proceed by direct calculation:

$$\begin{aligned}
(\Pi_\lambda \psi)(\mathbf{x}_{\sigma(1)}, \ldots, \mathbf{x}_{\sigma(n)}) &= \frac{1}{n!} \sum_{\pi \in S_n} \lambda(\pi^{-1}) \psi(\mathbf{x}_{(\pi\circ\sigma)(1)}, \ldots, \mathbf{x}_{(\pi\circ\sigma)(n)}) \,, \\
&= \frac{1}{n!} \sum_{\pi \in S_n} \lambda(\sigma \circ (\pi \circ \sigma)^{-1}) \psi(\mathbf{x}_{(\pi\circ\sigma)(1)}, \ldots, \mathbf{x}_{(\pi\circ\sigma)(n)}) \,, \\
&= \frac{1}{n!} \sum_{\tilde{\pi} \in S_n} \lambda(\sigma \circ \tilde{\pi}^{-1}) \psi(\mathbf{x}_{\tilde{\pi}(1)}, \ldots, \mathbf{x}_{\tilde{\pi}(n)}) \,, \\
&= \lambda(\sigma)(\Pi_\lambda \psi)(\mathbf{x}_1, \ldots, \mathbf{x}_n) \,.
\end{aligned}$$

To go from the second to the third line, we have used that for fixed $\sigma \in S_n$, as π ranges over S_n, so does $\tilde{\pi} = \pi \circ \sigma$ and so we can replace the latter by the former in the summation.

---

[34] Since for our multiplicative characters $\lambda(\pi) = \pm 1 = \lambda(\pi^{-1})$, the π^{-1} argument could be replaced with a π. The expression here is the one that generalises to more general finite groups with multiplicative characters that realise more general phase values.

Using this, we then confirm that

$$\begin{aligned}
\Pi_\lambda(\Pi_\lambda\psi)(\mathbf{x}_1,\dots,\mathbf{x}_n) &= \frac{1}{n!}\sum_{\pi\in S_n}\lambda(\pi^{-1})(\Pi_\lambda\psi)(\mathbf{x}_{\pi(1)},\dots,\mathbf{x}_{\pi(n)})\ , \\
&= \frac{1}{n!}\sum_{\pi\in S_n}\lambda(\pi^{-1})\lambda(\pi)(\Pi_\lambda\psi)(\mathbf{x}_1,\dots,\mathbf{x}_n)\ , \\
&= (\Pi_\lambda\psi)(\mathbf{x}_1,\dots,\mathbf{x}_n)\ ,
\end{aligned}$$

where we have used that $|S_n| = n!$. Finally, self-adjointness can be shown by term-by-term change of variables in the inner product. We leave the details to the interested reader. The stated result then follows. ■

4.2.1 Two-particle projections

For the two-particle case ($n = 2$), where the space of distinguishable-particle wave functions is $L^2(\mathbb{R}^d \times \mathbb{R}^d)$, the two projectors we have just defined are just the operations of taking symmetric and antisymmetric combinations, respectively:

$$\Pi_1\psi(\mathbf{x}_1,\mathbf{x}_2) = \frac{\psi(\mathbf{x}_1,\mathbf{x}_2)+\psi(\mathbf{x}_2,\mathbf{x}_1)}{2}\ , \qquad \Pi_\varepsilon\psi(\mathbf{x}_1,\mathbf{x}_2) = \frac{\psi(\mathbf{x}_1,\mathbf{x}_2)-\psi(\mathbf{x}_1,\mathbf{x}_2)}{2}\ . \tag{4.11}$$

In this case, all wave functions can be decomposed into symmetric and antisymmetric parts, so the full space of (distinguishable) two-particle wave functions can be decomposed into bosonic and fermionic wave functions. Alternatively, this can be phrased as the identity

$$\Pi_1 + \Pi_\varepsilon = 1_{L^2(\mathbb{R}^d\times\mathbb{R}^d)}\ , \tag{4.12}$$

which can be re-interpreted as the resolution of the identity for the permutation operator that exchanges $\mathbf{x}_1 \leftrightarrow \mathbf{x}_2$.

Note that this is not the situation for larger values of n; there are wavefunctions that cannot be decomposed into just totally-symmetric and totally-antisymmetric parts. We will make a related observation when we count bosonic and fermionic states associated to finite-dimensional Hilbert spaces later in this Chapter.

4.2.2 n-particle projections

Though the general projection operator is a little complicated to perform in practice for general wave functions (it involves choosing a sufficiently efficient way to sum over permutations), there is a case where things can be phrased more compactly. This is where we start with a *separable* distinguishable-particle wave function:

$$\psi(\mathbf{x}_1,\dots,\mathbf{x}_n) = \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\dots\psi_n(\mathbf{x}_n)\ . \tag{4.13}$$

This is a particularly natural class of wavefunctions to consider when considering *non-interacting* identical particles, where we might choose the ψ_i to be stationary states of the one-particle Hamiltonian acting on $\mathbf{x}_i$ to get stationary states for the full n-particle system.

The fermionic projection can then be realised in terms of what's called the *Slater determinant* of the single-particle wave-functions.

Definition 4.2.2. The *Slater determinant* of the wave-functions $\{\psi_i(\mathbf{x})\}$ is the determinants

$$|\psi_1,\dots,\psi_n\rangle = \frac{1}{\sqrt{n!}}\begin{vmatrix}
\psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \dots & \dots & \psi_n(\mathbf{x}_1) \\
\psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \dots & \dots & \psi_n(\mathbf{x}_2) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\psi_1(\mathbf{x}_n) & \dots & \dots & \dots & \psi_n(\mathbf{x}_n)
\end{vmatrix}\ . \tag{4.14}$$

The normalisation is such that if the ψ_i are mutually orthonormal, then $|\psi_1,\dots,\psi_n\rangle$ is normalised. We then have that

$$\Pi_\varepsilon\left(\psi_1(\mathbf{x}_1)\cdots\psi_n(\mathbf{x}_n)\right) = \frac{1}{\sqrt{n!}}|\psi_1,\dots,\psi_n\rangle\ . \tag{4.15}$$

Since arbitrary states can be expressed as (infinite) linear combinations of separable states, all fermionic wave functions can be obtained as (infinite) linear combinations of these types of states.

An analogous construction works in the bosonic case, though this uses the so-called *permanent* of a matrix, which is like the determinant but without the signs,

$$\Pi_1\psi = \frac{1}{n!}\,\mathrm{perm}\begin{pmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \dots & \dots & \psi_n(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \dots & \dots & \psi_n(\mathbf{x}_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \psi_1(\mathbf{x}_n) & \dots & \dots & \dots & \psi_n(\mathbf{x}_n) \end{pmatrix} . \tag{4.16}$$

As in the fermionic case, arbitrary bosonic wave functions can be constructed from these permanent states.

4.3 Symmetric and anti-symmetric tensor products

Though above we focused on fermionic and bosonic *wave functions*, the (anti-)symmetrisation procedure we have developed applies equally well to the case when we are taking tensor powers of some general Hilbert space $\mathcal{H}$ such as the qubit Hilbert space (or even a general vector space, for that matter). Here the n-fold tensor product of $\mathcal{H}$ admits a natural action of the symmetric group S_n just as was the case for wave functions: for $\psi_i \in \mathcal{H}$, we have

$$\begin{aligned} (S_n, \mathcal{H}^{\otimes n}) &\longrightarrow \mathcal{H}^{\otimes n} \\ \left(\pi, \psi_1 \otimes \cdots \otimes \psi_n\right) &\longmapsto \psi_{\pi(1)} \otimes \cdots \otimes \psi_{\pi(n)} . \end{aligned} \tag{4.17}$$

This action on pure tensors extends by linearity to all of $\mathcal{H}^{\otimes n}$.

We can then define bosonic and fermionic projection operators analogous to the ones we used for wave functions above. Just like we had for separable wave functions, we can define the action of these projection operators on pure tensors in $\mathcal{H}^{\otimes n}$ in terms of the determinants and permanents. We can analogously define bosonic and fermionic states in the n-fold tensor product of identical Hilbert spaces as the ranges of the corresponding orthogonal projectors:

Definition 4.3.1. The *n-fold symmetric tensor product* $\odot^n\mathcal{H}$ of the Hilbert space $\mathcal{H}$ is the subspace of the n-fold tensor product $\mathcal{H}^{\otimes n}$ on which Π_1 acts as the identity, or equivalently,

$$\odot^n\mathcal{H} = \mathrm{Ran}_{\mathcal{H}^{\otimes n}}\Pi_1 . \tag{4.18}$$

This is sometimes also denoted $\mathrm{Sym}^n\mathcal{H}$, and these are states that are compatible with Bose–Einstein statistics.

Definition 4.3.2. The *n-fold antisymmetric tensor product* of the Hilbert space $\mathcal{H}$ is the subspace of the n-fold tensor product $\mathcal{H}^{\otimes n}$ on which Π_ε acts as the identity, or alternatively,

$$\wedge^n\mathcal{H} = \mathrm{Ran}_{\mathcal{H}^{\otimes n}}\Pi_\varepsilon . \tag{4.19}$$

This is sometimes called the *exterior tensor product*, and these states are compatible with Fermi–Dirac statistics.

From the standard properties of determinants, a Slater determinant state will vanish identically if two of the constituent ψ_i are proportional. This means that the basis of n-particle states we get by acting with the fermionic projection operator on a basis of pure tensors all come from states where each of the n particles is in a distinct basis state. This is often phrased in terms of the following,

The Pauli exclusion principle: *Two fermions cannot occupy the same state.*

Indeed, this leads to a significant reduction in the number of fermionic states that can be constructed from a given set of single-particle states. To see this more explicitly, let us count the bosonic and fermionic states that can be built from a given N-dimensional Hilbert space under iterated symmetric and anti-symmetric tensor products.

Lemma 4.3.3. The space of fermionic n-particle states built from an N-dimensional single-particle Hilbert space $\mathcal{H}$ has dimension given by

$$\dim\left(\wedge^n\mathcal{H}\right) = \binom{N}{n} \ . \tag{4.20}$$

Proof. We choose a basis for $\mathcal{H}$ and build a basis of states for the fermionic n-particle Hilbert space using Slater determinants where the ψ_i are elements of that basis. There are N choices for ψ_1, but since ψ_2 cannot be the same as ψ_1, there are $N-1$ choices for ψ_2 and so on. The final state is independent of the ordering of $\psi_1, \ldots, \psi_n$ so we have

$$\dim\left(\wedge^n\mathcal{H}\right) = \frac{N(N-1)\cdots(N-n+1)}{n!} = \frac{N!}{n!(N-n)!} = \binom{N}{n} \ . \tag{4.21}$$

■

Which leads to an immediate important observation, which will play an important role when we consider atomic structure in the presence of several electrons.

Corollary 4.3.4. At most N identical, non-interacting fermionic particles can coexist in a given N-dimensional single-particle Hilbert space $\mathcal{H}$.

For completeness, we also consider bosonic multi-particle states built from a given N-dimensional Hilbert space.

Lemma 4.3.5. The space of bosonic n-particle states built from an N-dimensional single-particle Hilbert space $\mathcal{H}$ has dimension given by

$$\dim\left(\odot^n\mathcal{H}\right) = \frac{(N+n-1)!}{(N-1)!\,n!} \ . \tag{4.22}$$

Proof. In order to prove this we introduce a generating function known as a *partition function* that has much wider applicability. In general a separable bosonic state can be represented as

$$\Pi_1\left(\psi_1^{\otimes k_1}\otimes\psi_2^{\otimes k_2}\cdots\psi_N^{\otimes k_N}\right) \ , \qquad \sum k_i = n \ . \tag{4.23}$$

The overall order doesn't matter because of the symmetrisation, so we only pay attention to how many times each basis element appears. We therefore want to count the number of non-negative integer partitions of n, ($\{k_i \in \mathbb{Z}_{\geqslant 0}\}$ such that the $\sum_i k_i = n$). Let us replace the ψ_i by formal variables x_i, whereupon our problem becomes that of counting the number of distinct monomials of the form $x_1^{k_1}x_2^{k_2}\cdots x_N^{k_N}$ of total degree n. If we further multiply each x_i by an additional formal variable s, then the total power of s will be the total degree. Taking the sum over *all* $k_1, \ldots, k_N$, we obtain

$$\sum_{k_1,\ldots,k_N\in\mathbb{N}} (sx_1)^{k_1}(sx_2)^{k_2}\ldots(sx_N)^{k_N} = \prod_{j=1}^{N}\sum_{k_j=0}^{\infty}(sx_j)^{k_j} = \prod_{j=1}^{N}\frac{1}{1-sx_j} \ . \tag{4.24}$$

If we set all the x_i to one, we obtain $1/(1-s)^N$ and the coefficient of s^n will simply count the number of terms where $\sum k_i = n$. The generalised binomial theorem then gives the following, from which the result follows.

$$\frac{1}{(1-s)^N} = \sum_{n\in\mathbb{N}}\binom{N+n-1}{n}s^n \ , \tag{4.25}$$

■

One can now observe explicitly that while for $n = 2$ there is an accidental equality

$$\dim\mathcal{H}^{\otimes n} = N^n = \dim\wedge^n\mathcal{H} + \dim\odot^n\mathcal{H} \ , \qquad n = 2 \ , \tag{4.26}$$

for more than two particles we have

$$\dim\mathcal{H}^{\otimes n} = N^n > \dim\wedge^n\mathcal{H} + \dim\odot^n\mathcal{H} \ , \qquad n > 2 \ . \tag{4.27}$$

So for more than two particles, a general multi-particle state cannot be decomposed into bosonic and fermionic parts.

Chapter 5

Symmetries and Unitary Groups

In this chapter, we turn our attention to the realisation of *symmetries* in quantum theories/systems. Before developing a formal theory of quantum symmetry, we will look at a simple example that illustrates many features of the formalism and most of the key ideas.

5.1 An appetizer: spatial and time translations

Consider a particle moving freely (subject to no forces) on the real line (so, as in Chapter 1.3, the Hilbert space is identified as $\mathcal{H} \cong L^2(\mathbb{R})$, and additionally the Hamiltonian is just $H = P^2/2m$). There is a clear physical sense in which *linear translations* in space should be symmetries of the theory, as they are in the classical setting. How does this manifest in the quantum mechanical formalism?

For $a \in \mathbb{R}$, we can define an operator $T(a) : \mathcal{H} \to \mathcal{H}$ that corresponds to a translation of the entire system by a relative to a fixed reference frame. This will act on wavefunctions according to

$$(T(a)\psi)(x) = \psi(x-a) \ . \tag{5.1}$$

To understand why the minus sign, observe that the value of the transformed wavefunction at a will be the value of the original wave function at the origin. In terms of our generalised position eigenstates, we have[35]

$$T(a)\,|\xi\rangle = |\xi + a\rangle \ , \tag{5.2}$$

because a (generalised) eigenstate that was previously localised at $x = \xi$ should be localised at $x = \xi + a$ after translation. We see this is equivalent to (5.1) as follows,

$$\begin{aligned}
(T(a)\psi)(x) = \langle x|T(a)|\psi\rangle &= \int_{-\infty}^{\infty} \mathrm{d}\xi \ \langle x|T(a)|\xi\rangle\,\langle\xi|\psi\rangle \ , \\
&= \int_{-\infty}^{\infty} \mathrm{d}\xi \ \langle x|\xi + a\rangle\,\psi(\xi) \ , \\
&= \int_{-\infty}^{\infty} \mathrm{d}\xi\,\delta(x - \xi - a)\psi(\xi) \ , \\
&= \psi(x-a) \ .
\end{aligned} \tag{5.3}$$

We make some immediate observations regarding the structural properties of these translation operators:

(1) $T(a)T(b) = T(a+b) \quad \forall\, a, b \in \mathbb{R}\,,$

(2) $T(a)^{-1} = T(-a) \qquad \forall\, a \in \mathbb{R}\,,$

(3) $T(0) = 1\,.$

The adjoint of this translation operator is determined by a change of variables in the integral expression for the inner product,

$$\langle\chi|T(a)\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\chi(x)\psi(x-a) = \int_{-\infty}^{\infty} \mathrm{d}x\,\chi(x+a)\psi(x) = \langle T(-a)\chi|\psi\rangle \ . \tag{5.4}$$

so we have

---

[35] Keeping track of the signs here is a good exercise in disambiguating generalised position eigenstates from their wavefunctions

(4) $T(a)^* = T(-a) = T(a)^{-1}$

As in our discussion of time translation, this last condition identifies the operators $T(a)$ as unitary operators. The four properties that we have listed then precisely identify this structure as a *unitary representation of the additive group* $(\mathbb{R}, +)$ *on the Hilbert space* $\mathcal{H}$.

Definition 5.1.1. The *unitary group* $\mathsf{U}(\mathcal{H})$ is the group of unitary operators on the Hilbert space $\mathcal{H}$. For the case of when $\mathcal{H}$ is finite-dimensional (say $\dim(\mathcal{H}) = n$), this can be identified with the usual matrix group $U(n)$.

Definition 5.1.2. A *unitary representation* of a group $\mathcal{G}$ on a Hilbert space $\mathcal{H}$ is a group homomorphism from $\mathcal{G}$ to $\mathsf{U}(\mathcal{H})$.

In the case of infinite-dimensional $\mathcal{H}$, such a group homomomorphism is required to be what is called *strongly continuous*. We will not need to pay attention to that restriction in this course and so will not define it carefully—what it means is roughly that for any $\psi \in \mathcal{H}$, the induced map from $\mathcal{G}$ to $\mathcal{H}$ is continuous. (See, though, Def. 5.2.5 below for a related condition.)

An important feature of the group of translations is that they can be taken arbitrarily small, in which case the translation operator should become (in a suitable sense) arbitrarily close to the identity operator. We can observe how this transpires in terms of the action of translations on (differentiable) wave functions:[36]

$$\lim_{\varepsilon\to 0}(T(\varepsilon)\psi)(x) = \lim_{\varepsilon\to 0}\psi(x-\varepsilon) = \psi(x) - \varepsilon\psi'(x) + O(\varepsilon^2)\ . \tag{5.5}$$

We interpret this result as defining an infinitesimal expansion of the translation operator itself,

$$T(\varepsilon) = 1 - \frac{i\varepsilon}{\hbar}T_{\text{inf}} + O(\varepsilon^2)\ , \qquad (T_{\text{inf}}\psi)(x) = -i\hbar\psi'(x)\ . \tag{5.6}$$

We have inserted conventional factors of i and $\hbar$ that allows us to make the identification of T_{inf} with the *momentum operator* P.

The relationship between translations and momentum is easier to tease out in momentum space. On generalised momentum eigenstates, we have

$$\begin{aligned} T(a)\,|p\rangle &= \int_{-\infty}^{\infty} \mathrm{d}x\, T(a)\,|x\rangle\,\langle x|p\rangle &&= \int_{-\infty}^{\infty} \mathrm{d}x\,|x+a\rangle\,\mathcal{N}e^{\frac{ipx}{\hbar}}\ , \\ &= \int_{-\infty}^{\infty} \mathrm{d}x\,|x\rangle\,\mathcal{N}e^{\frac{ip(x-a)}{\hbar}} &&= e^{-\frac{ipa}{\hbar}}\int_{-\infty}^{\infty} \mathrm{d}x\,|x\rangle\,\mathcal{N}e^{\frac{ipx}{\hbar}}\ , \\ &= e^{-\frac{ipa}{\hbar}}\,|p\rangle &&= e^{-\frac{iPa}{\hbar}}\,|p\rangle\ , \end{aligned} \tag{5.7}$$

so on our (continuum) basis of generalised momentum eigenstates we have the operator relation

$$T(a) = \exp\left(-\frac{iPa}{\hbar}\right)\ , \tag{5.8}$$

from which we formally deduce $T_{inf} = P$ just by taking the power series expansion of the exponential. We summarise this situation by saying that P is the *infinitesimal generator of translations*.[37] Note also that unitarity of $T(a)$ follows

---

[36] The restriction to differentiable wave functions here is, once again, related to the infinite dimensionality of our Hilbert space, which means that the infinitesimal version of translation that we are defining is only partially defined on $L^2(\mathbb{R})$. Differentiable wave functions are dense in $L^2(\mathbb{R})$.

[37] If you are familiar with Noether's theorem from classical mechanics, then this should sound familiar as a counterpart of the fact that momentum is the conserved quantity associated with translation invariance, and it generates infinitesimal translations via the Poisson bracket.

from self-adjointness of P and vice versa,

$$T(a)^* = \left(e^{-\frac{iPa}{\hbar}}\right)^* = \left(e^{\frac{iPa}{\hbar}}\right) = T(a)^{-1} . \tag{5.9}$$

Thus far this discussion makes no reference to translations being a *dynamical symmetry* of the system, *i.e.*, being a symmetry of the equations of motion (in our case, the time-dependent Schrödinger equation). Physically, this should depend on the potential $V(X)$ being constant (or zero), as otherwise the potential would violate translation invariance. To have translations as a dynamical symmetry, we would like to require that the symmetry transformation of the infinitesimal time evolution of a state vector is the same as the infinitesimal time evolution of the symmetry-transformed state vector, *i.e.*,

$$T(a)\,|H\psi\rangle = H\,|T(a)\psi\rangle \ , \tag{5.10}$$

which, by writing (without loss of generality) $|\psi\rangle = T(a)^*\,|\phi\rangle$ for some $|\phi\rangle$, we can equivalently characterise as

$$T(a)HT(a)^* = H . \tag{5.11}$$

Further looking at the case of infinitesimal translations, this gives the condition

$$\left(1 - \frac{i\varepsilon}{\hbar}P + O(\varepsilon^2)\right) H \left(1 + \frac{i\varepsilon}{\hbar}P + O(\varepsilon^2)\right) = H + \frac{i\varepsilon}{\hbar}[H, P] + O(\varepsilon^2) = H , \tag{5.12}$$

from which we deduce the requirement $[H, P] = 0$. Using our expression for finite translations as an exponentiated version of P, one can show that this vanishing commutator also implies the relation (5.11). And indeed, these will hold for a Hamiltonian of the form $H = P^2/2m + V(X)$ only if $V(X)$ is a constant.[38]

5.2 A more general theory of quantum symmetries

What we have seen above gives us some insight into the mathematical representation of of symmetries in quantum systems. Now we will look at this topic more abstractly.

5.2.1 A first attempt at generalisation

If we try to generalise a bit from what we have seen in our example, we might propose the following structures associated with the presence of a symmetry in a quantum system. It will turn out that these are not quite the complete story; we will return to the correct formulation after some technical discussion.

- Symmetries should be implemented via unitary operators on $\mathcal{H}$, so as to preserve norms and transition amplitudes.
- Symmetries naturally form a group—call it $\mathcal{G}$—and the operators implementing their action on $\mathcal{H}$ should form a unitary representation of that group,

$$U : \mathcal{G} \to \mathsf{U}(\mathcal{H}) \ , \qquad U(g_1)U(g_2) = U(g_1g_2) \quad \forall\, g_1, g_2 \in \mathcal{G} . \tag{5.13}$$

- For continuous symmetries, infinitesimal transformations are realised by self-adjoint operators that generate finite transformations (parameterised by $s \in \mathbb{R}$) via exponentiation according to

$$U(g(s)) = \exp\left(-\frac{iGs}{\hbar}\right) \ , \quad G = G^* . \tag{5.14}$$

- For dynamical symmetries, we require

$$U(g)HU(g^{-1}) = H , \qquad ([H, G] = 0 \quad \text{for infinitesimal generators}) . \tag{5.15}$$

---

[38] As a concrete example of this, in your first homework exercise, you will have shown that for the harmonic oscillator the finite spatial translation of the ground state is a coherent state, which is certainly no longer an energy eigenstate, let alone the ground state.

These properties do hold in quite a few examples of interest. For instance, you should compare the above to our discussion of time evolution in Chapter 1. However, they are not the most general version of the story, and we have also been a bit cavalier about some technical details in our discussion of infinitesimal symmetries. We address both of these issues below.

5.2.2 Quantum symmetries and projective representations

The main shortcoming of the formulation above arises from having neglected the distinction between Hilbert space $\mathcal{H}$ and the true space of quantum states, $\mathbb{P}(\mathcal{H})$. *A priori*, one expects that a quantum symmetry need only be formulated as a map

$$\mathfrak{s}:\mathbb{P}(\mathcal{H})\longrightarrow\mathbb{P}(\mathcal{H})\,. \tag{5.16}$$

Rather than requiring that overlaps be preserved, it should be sufficient to require that *transition probabilities* are preserved, as these are the physically meaningful quantities. Let us denote a quantum state corresponding to the ray in $\mathcal{H}$ that passes through a vector ψ by $[\psi]$, so $[\psi]=[\lambda\psi]$ for $\lambda\in\mathbb{C}^\times$. For quantum states $[\psi],[\varphi]\in\mathbb{P}(\mathcal{H})$, we then require equality of the transition probabilities:

$$\frac{|\langle\varphi|\psi\rangle|^2}{\|\varphi\|^2\|\psi\|^2}=\frac{|\langle\mathfrak{s}(\varphi)|\mathfrak{s}(\psi)\rangle|^2}{\|\mathfrak{s}(\varphi)\|^2\|\mathfrak{s}(\psi)\|^2}\,, \tag{5.17}$$

where in this expression, ψ and φ could be any representatives of the quantum states $[\psi]$ and $[\varphi]$, respectively. (The transition probabilities are, as usual, independent of the choice of such representative.) Naively, it appears that this could be a weaker condition than the requirement of a unitary map on $\mathcal{H}$. This is indeed the case, but perhaps to a lesser extent than one might first think. The situation is explained by the following.

Theorem 5.2.1 (Wigner)**.** For any quantum symmetry $\mathfrak{s}$ defined as above on projectivised Hilbert space, there exists an operator $V(\mathfrak{s}):\mathcal{H}\to\mathcal{H}$ that is compatible with $\mathfrak{s}$ that is either unitary or anti-unitary that induces $\mathfrak{s}$ when treated as a map of rays. When $\dim(\mathcal{H})\geqslant 2$, the operator $V(\mathfrak{s})$ is unique up to an overall phase.

(In the case that $\dim(\mathcal{H})=1$, $V(\mathfrak{s})$ can be chosen to be *either* unitary or anti-unitary for the same $\mathfrak{s}$; in higher dimensional Hilbert spaces it will be one or the other, with no choice involved other than the aforementioned phase.) We will set aside the topic of anti-unitary operators for the moment and focus on symmetries that are realised as unitary operators on $\mathcal{H}$.

Definition 5.2.2. The *projective unitary group* of a Hilbert space $\mathcal{H}$ is the quotient

$$\mathbb{P}\mathsf{U}(\mathcal{H})=\mathsf{U}(\mathcal{H})/\{e^{i\theta}1_{\mathcal{H}},\ \theta\in\mathbb{R}\}$$

of the group of unitary transformations on $\mathcal{H}$ by the normal subgroup consisting of multiplications by a constant phase.

What Wigner's theorem is telling us, given this definition, is that (neglecting the anti-unitary caveat) a quantum symmetry can be unambiguously lifted to an element of the projective unitary group for the corresponding Hilbert space.

Definition 5.2.3. A *projective unitary representation* of a group $\mathcal{G}$ on a Hilbert space $\mathcal{H}$ is a group homomorphism $U:\mathcal{G}\to\mathbb{P}\mathsf{U}(\mathcal{H})$.

What we should then be interested in are these projective unitary representations of a symmetry group $\mathcal{G}$. We can then lift these symmetries to actual unitary operators, but we have to choose phases. If we do this arbitrarily (choosing arbitrary phases for each $g\in\mathcal{G}$), then at the level of unitary operators the group law may not be obeyed, indeed we only expect

$$U(g_1)U(g_2)=e^{i\xi(g_1,g_2)}U(g_1g_2)\,,\qquad \xi(g_1,g_2)\in[0,2\pi)\,. \tag{5.18}$$

Associativity of the multiplication of operators on $\mathcal{G}$ gives a condition on these phases,[39]

$$\xi(g_1,g_2g_3)+\xi(g_2,g_3)=\xi(g_1,g_2)+\xi(g_1g_2,g_3)\mod 2\pi\,, \tag{5.19}$$

---

[39] Though it is not important for us, this condition means that the map $\xi:\mathcal{G}\times\mathcal{G}\to U(1)$ is what is known as a *group 2-cocycle valued in* $U(1)$.

It follows immediately that we have the following

Proposition 5.2.4. A *projective unitary representation* of a group $\mathcal{G}$ on a Hilbert space $\mathcal{H}$ is equivalently a map $U : \mathcal{G} \to \mathrm{U}(\mathcal{H})$ obeying (5.18) and (5.19).

Indeed, this is sometimes used as the definition; physicists will say that "in quantum mechanics, the group law for symmetries only needs to be obeyed up to phase ambiguities". The phase ambiguities in the definitions of the operators $U(g_i)$ themselves means we can modify the phases in (5.18) by taking a map $\phi : \mathcal{G} \to \mathbb{S}^1$ to produce an equivalent projective unitary representation but now with

$$\xi(g_1, g_2) \to \xi(g_1, g_2) + \phi(g_1) + \phi(g_2) \mod 2\pi \,. \tag{5.20}$$

Using this freedom to redefine phases, it turns out that in a large class of examples of continuous groups (when either $\mathcal{H}$ is finite-dimensional or, if infinite-dimensional, if $\mathcal{G}$ is what is called a *semi-simple* group), one can set the phases $\xi(g_1, g_2)$ to be zero for transformations that are suitably close to the identity, thus realising something that looks like a unitary representation for symmetries that are close to the identity. Globally there can be an obstruction to setting these phases to zero for all group multiplications; we will see an explicit example of this in the next chapter in the context of rotations.

5.2.3 One parameter unitary groups

In the discussion of translations some of the statements about the infinitesimal limit may have felt a bit sketchy. There is actually a powerful theorem that puts these statements on firm footing. We start with a definition:

Definition 5.2.5. A *strongly continuous one-parameter unitary group* is a family $U(t)$ for $t \in \mathbb{R}$ of unitary operators on a Hilbert space $\mathcal{H}$ such that

- $U(0) = 1_{\mathcal{H}}$,
- $\forall s, t \in \mathbb{R}$, $\quad U(t + s) = U(t)U(s)$,
- $\forall t \in \mathbb{R}$, $\quad \lim_{s \to t} U(s)\psi = U(t)\psi$.

The first two points defines a one-parameter unitary group, which you will recognise as being the same as a unitary representation of the additive group $(\mathbb{R}, +)$ as we had in the case of translations. The third point is the notion of *strong continuity*. We will not be very attentive to this continuity condition in this course; it will always hold in examples we consider.

Definition 5.2.6. For $U(\cdot)$ a strongly continuous one-parameter unitary group, the *infinitesimal generator* of $U(\cdot)$ is the operator K defined by[40]

$$K\psi = \lim_{t \to 0} \frac{1}{i} \frac{U(t)\psi - \psi}{t} \,. \tag{5.21}$$

It turns out with this definition, K will be defined for a dense subset of $\mathcal{H}$ (or all of $\mathcal{H}$ in the finite-dimensional case). We then have the following:

Theorem 5.2.7 (Stone's Theorem on One-Parameter Unitary Groups)**.** Let $U(\cdot)$ be a strongly continuous, one-parameter unitary group. The infinitesimal generator K of the family is a self-adjoint operator, and for all t we have,

$$U(t) = \exp(itK) \,. \tag{5.22}$$

Conversely, every self-adjoint operator K generates a strongly continuous one-parameter unitary group this way.

---

[40] This definition differs by a conventional minus sign and factor of $\hbar$ relative to what we used in the case of momentum and translations. In the quantum mechanical setting we will normally include those additional factors.

The exponential of K can be defined in terms of its action on a basis of (generalised) K-eigenstates. We will not study a proof of this theorem; instead it is meant to provide a justification for some of the more casual manipulations that have arisen and will arise when studying symmetries in what follows. It should be noted that even in the finite dimensional case this is a non-trivial theorem; it establishes a kind of differentiability for families of operators/matrices based only upon continuity.

We can observe now that the quantity $U(t_1 - t_0)$ that we considered in Chapter 1 and in the discussion of the propagator is precisely a one-parameter unitary group whose infinitesimal generator is the Hamiltonian. This is in the case when the Hamiltonian is time-independent. In the time-dependent case, one actually gets a unitary *groupoid*! We will not discuss this further.

5.2.4 Anti-unitary operators

Now we return to the issue of anti-unitary operators, which appeared in the statement of Wigner's theorem, and define such an entity.

Definition 5.2.8. An *anti-unitary* operator on a Hilbert space $\mathcal{H}$ is a surjective linear map $A : \mathcal{H} \to \mathcal{H}$ obeying

$$\langle A\varphi | A\psi \rangle = \overline{\langle \varphi | \psi \rangle} = \langle \psi | \varphi \rangle \ . \tag{5.23}$$

We can see that an anti-unitary operator must be $\mathbb{C}$ anti-linear. A standard example of an anti-unitary operator on a complex Hilbert space is a complex conjugation operation, which takes states of the form

$$\psi = \sum_i c_i \psi_i \quad \longrightarrow \quad A\psi = \sum_i \overline{c_i} \psi_i \ , \tag{5.24}$$

for $\{\psi_i\}$ an orthonormal basis. (This operation clearly depends on the basis.) In the case of $L^2(\mathbb{R})$, one has a similar operation that takes the complex conjugate of a wave function.

An important observation is that if A is anti-unitary, then A^2 is unitary,

$$\langle A^2\varphi | A^2\psi \rangle = \langle A\psi | A\varphi \rangle = \langle \varphi | \psi \rangle \ . \tag{5.25}$$

This means that any symmetry that can be realised as the square of another symmetry will be realised unitarily on $\mathcal{H}$. In the case of continuous groups of symmetries, like translations and rotations, this lets us get away with ignoring anti-unitary symmetries all together. On a homework exercise, you will investigate the relationship between anti-unitary symmetries and time-reversal.

5.2.5 The form of quantum symmetries

We are now in position to formulate a more precise characterisation of the form that symmetries take in quantum mechanical systems.

- Symmetries are implemented via unitary or anti-unitary operators on $\mathcal{H}$.
- Symmetries naturally form a group, and the operators implementing them form a projective representation of that group on $\mathcal{H}$.
- Continuous symmetries are generated, in the sense of Stone's theorem, by self-adjoint operators via exponentiation.
- For unitarily realised symmetries to be compatible with time evolution (dynamical symmetries), we require any of the following equivalent conditions
  - $U(t)U(g) = U(g)U(t)$,
  - $[H, U(g)] = 0$

  - $[U(t), G] = 0$
  - $[H, G] = 0$

Here $U(t) = \exp(-iHt/\hbar)$ is the time evolution operator while $U(g)$ is the unitary corresponding to an element $g \in \mathcal{G}$ of the symmetry group. If $U(g)$ is part of a one-parameter group, then its infinitesimal generator is G.

Chapter 6

Rotations, Angular Momentum, and Their Representations

In this chapter, we analyse a crucial symmetry that appears time and again in important quantum systems: that of three-dimensional rotations. You know well from your geometry course (and perhaps elsewhere) that the proper rotation group in three dimensions is $\mathrm{SO}(3)$, which can be identified with the group of three-by-three orthogonal matrices with unit determinant. In line with the general structures described in the previous chapter, we expect that for a quantum system describing objects in three-dimensions, there should be a (projective) unitary representation of $\mathrm{SO}(3)$ on our Hilbert space. We will have seen how this cashes out in practice by the end of the chapter.

6.1 Rotation group $\mathrm{SO}(3)$ and its infinitesimal generators

We begin with a review of some technical aspects of the the three-dimensional orthogonal group $\mathrm{O}(3)$. This group is normally realised as a group of three-by-three real matrices acting on Cartesian coordinates $x_i = (x_1, x_2, x_3)$ according to,

$$x_i \longrightarrow \sum_{j=1}^{3} R_{ij} x_j \,, \qquad RR^{\mathsf{T}} = \mathbb{1}_{3\times 3} \,, \tag{6.1}$$

where $\mathbb{1}_{3\times 3}$ is the three-by-three identity matrix. The special orthogonal group $\mathrm{SO}(3)$ restricts to those transformations that are rotations—it is the subgroup of $\mathrm{O}(3)$ for which $\det(R) = 1$.

As with translations, rotations can be taken arbitrarily close to the identity. To characterise this, let us consider a one-parameter family of rotation matrices $R(t)$ with $R(0) = \mathbb{1}_{3\times 3}$. (You may wish to think of this as the family of rotations about a fixed axis with t proportional to the angle of rotation.) We can define the matrix elements of an infinitesimal rotation matrix ω according to

$$R_{ij}(t) = \delta_{ij} + t\omega_{ij} + O(t^2) \,. \tag{6.2}$$

Expanding the condition $R(t)R^{\mathsf{T}}(t) = \mathbb{1}_{3\times 3}$ to first order in t (or alternatively differentiating at $t = 0$) gives

$$\omega_{ij} + \omega_{ji} = 0 \,, \tag{6.3}$$

or in matrix notation, $\omega + \omega^{\mathsf{T}} = 0$, *i.e.*, ω is a *skew symmetric* matrix. As you saw in prelims, it is natural to organise the components ω_{ij} of this matrix into a vector $\omega = (\omega_1, \omega_2, \omega_3) = (\omega_{32}, \omega_{13}, \omega_{21})$, and this encodes the axis about which the instantaneous rotation is taking place and the rate of rotation (in the vector's magnitude). The vector and matrix index labelling for these parameters are related according to

$$\omega_i = -\frac{1}{2}\sum_{j,k} \varepsilon_{ijk}\omega_{jk} \,, \qquad \omega_{ij} = -\sum_{k} \varepsilon_{ijk}\omega_k \,. \tag{6.4}$$

The first-order action of $R(t)$ on the coordinate x_i then is given by

$$\begin{aligned} R(t)\mathbf{x} &= \mathbf{x} + t\delta\mathbf{x} + O(t^2) \,, \\ \delta x_i &= \sum_j \omega_{ij} x_j = \sum_j \varepsilon_{ijk}\omega_j x_k = (\omega \wedge \mathbf{x})_i \,. \end{aligned} \tag{6.5}$$

The group $\mathrm{SO}(3)$ is non-Abelian, so in general pairs of rotations do not commute, *i.e.*, $R\tilde{R} \neq \tilde{R}R$. This lack of commutativity is encoded in the *commutator* $R\tilde{R}R^{-1}\tilde{R}^{-1}$, which is itself element of $\mathrm{SO}(3)$ that will be the identity if and only if R and $\tilde{R}$ commute. Let us consider this commutator at the level of infinitesimal rotations. If we take t small in $R(t)$

and $\tilde{R}(t)$, then expanding the commutator to second order we have[41]

$$\begin{aligned} R(t)\tilde{R}(t)R(t)^{-1}\tilde{R}(t)^{-1} &= (1+t\omega+\ldots)(1+t\tilde{\omega}+\ldots)(1-t\omega+\ldots)(1-t\tilde{\omega}+\ldots)\ , \\ &= 1+t^2(\omega\tilde{\omega}-\tilde{\omega}\omega)+\ldots\ , \end{aligned} \tag{6.6}$$

so here the noncommutativity manifests in terms of the matrix commutator $[\omega,\tilde{\omega}]=\omega\tilde{\omega}-\tilde{\omega}\omega$. Notice that

$$[\omega,\tilde{\omega}]_{ik}=\sum_j \omega_{ij}\tilde{\omega}_{jk}-\tilde{\omega}_{ij}\omega_{jk}=-\sum_l \varepsilon_{ikl}(\omega\wedge\tilde{\omega})_l\ , \tag{6.7}$$

where to prove this it is useful to use the identity $\sum_k \varepsilon_{ijk}\varepsilon_{klm}=\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl}$. Alternatively, if we tacitly use (6.4) to identify vectors with skew-symmetric matrices,

$$[\omega,\tilde{\omega}]=(\omega\wedge\tilde{\omega})\ . \tag{6.8}$$

The vector space of three-by-three skew-symmetric matrices endowed with the bilinear operation of the matrix commutator (observe that this preserves skew-symmetry) is known as the *Lie algebra* $\mathfrak{so}(3)$. In the theory of Lie groups, one finds that this matrix commutator encodes the full structure of the group $\mathrm{SO}(3)$ up to a single ambiguity, to which we will return later in our discussion of spin.

6.2 Rotations and wave functions

As our first example, we can define a an action of the rotation group on wave functions in three dimensions, *i.e.*, on the Hilbert space $L^2(\mathbb{R}^3)$, in a natural manner:

$$\begin{aligned} \mathrm{SO}(3)\times L^2(\mathbb{R}^3) &\longrightarrow L^2(\mathbb{R}^3)\ , \\ (R,\psi) &\longmapsto (U(R)\psi)\ , \qquad \begin{aligned} (U(R)\psi)(R\mathbf{x}) &= \psi(\mathbf{x})\ , \\ (U(R)\psi)(\mathbf{x}) &= \psi(R^\intercal\mathbf{x})\ . \end{aligned} \end{aligned} \tag{6.9}$$

The appearance of the transpose (*i.e.*, inverse) in the argument is analogous to the minus sign that we included in our translation operator, and analogously to that case we have for generalised position eigenstates,

$$U(R)\,|\mathbf{x}\rangle=|R\mathbf{x}\rangle\ . \tag{6.10}$$

This action is manifestly complex linear. It is also unitary, since we have

$$\langle U(R)\psi|U(R)\psi\rangle=\int_{\mathbb{R}^3}|\psi(R^\intercal\mathbf{x})|^2\,\mathrm{d}^3\mathbf{x}=\int_{\mathbb{R}^3}|\psi(\tilde{\mathbf{x}})|^2\,\mathrm{d}^3\tilde{\mathbf{x}}=\langle\psi|\psi\rangle\ , \tag{6.11}$$

where the change of variables $\mathbf{x}\to\tilde{\mathbf{x}}=R^\intercal\,\mathbf{x}$ introduces no Jacobian because R is an orthogonal matrix. Under composition, we see the importance of the transpose:[42]

$$(U(R_1)U(R_2)\psi)(\mathbf{x})=(U(R_2)\psi)(R_1^\intercal\mathbf{x})=\psi(R_2^\intercal R_1^\intercal\mathbf{x})=\psi\left((R_1R_2)^\intercal\mathbf{x}\right)=(U(R_1R_2)\psi)(\mathbf{x})\ . \tag{6.12}$$

so our operators satisfy the group law,

$$U(R_1R_2)=U(R_1)U(R_2)\ , \tag{6.13}$$

and we have a unitary representation of $\mathrm{SO}(3)$.

Let us consider the infinitesimal version of this action. Using the expansion for rotation matrices in Equation (6.2), we

---

[41] You can feel free to take this equation for granted, but deriving it while keeping second-order terms might be instructive.

[42] You might try rewrite the manipulations in Equation (6.12) using bra-ket notations for wave functions to get a feeling for the way the compositions here are behaving and the relation to the action on generalised position eigenstates.

have

$$\begin{aligned}\psi(R^{\mathsf{T}}(t)\mathbf{x}) &= \psi\left(\mathbf{x} - t\omega\wedge\mathbf{x} + O(t^2)\right) ,\\ &\approx \psi(\mathbf{x}) - t(\omega\wedge\mathbf{x})\cdot\nabla\psi(\mathbf{x}) ,\\ &= \psi(\mathbf{x}) - t\,\omega\cdot(\mathbf{x}\wedge\nabla\psi) ,\\ &= \left(1_{L^2(\mathbb{R})} - \frac{it}{\hbar}\omega\cdot\mathbf{L}\right)\psi(\mathbf{x}) .\end{aligned} \tag{6.14}$$

where $\mathbf{L}$ is the *orbital angular momentum* operator that you met in **A11 Quantum Theory**, which we can rewrite in terms of position and momentum operators,

$$\mathbf{L} := \mathbf{X}\wedge\mathbf{P} . \tag{6.15}$$

You have seen in that previous course, and one can compute explicitly, that the components L_i of the angular momentum operator obey the commutation relations

$$[L_i, L_j] = i\hbar\sum_k \varepsilon_{ijk}L_k . \tag{6.16}$$

For general vectors ω and $\tilde{\omega}$, one then finds

$$[\omega\cdot\mathbf{L}, \tilde{\omega}\cdot\mathbf{L}] = i\hbar(\omega\wedge\tilde{\omega})\cdot\mathbf{L} , \tag{6.17}$$

or alternatively, in terms of the infinitesimal generators with extra constants included,

$$\left[-\frac{i}{\hbar}\omega\cdot\mathbf{L}, -\frac{i}{\hbar}\tilde{\omega}\cdot\mathbf{L}\right] = -\frac{i}{\hbar}(\omega\wedge\tilde{\omega})\cdot\mathbf{L} , \tag{6.18}$$

We observe that these exactly match the commutation relation (6.8) with the replacement

$$\omega \quad\longleftrightarrow\quad -\frac{i}{\hbar}\omega\cdot\mathbf{L} , \tag{6.19}$$

where on the left hand side, ω represents a skew-symmetric matrix, and on the right hand side ω is a vector indicating the axis of rotation and we have operators on $L^2(\mathbb{R})$. We say that these operators furnish a *representation of the Lie algebra* $\mathfrak{so}(3)$ *on the Hilbert space* $L^2(\mathbb{R})$.

6.3 General unitary representations

In the previous analysis, we had a manifest action of the rotation group on the space of wave functions. In a more general and abstract setting, we must consider a general (projective) unitary representation of the rotation group on a Hilbert space $\mathcal{H}$. This isn't such an easy thing to get our hands on, so we will approach the problem through infinitesimal rotations. We introduce infinitesimal generators of rotations (in the sense of Stone's theorem) and denote them by $\mathbf{J}$. For a one-parameter families of rotations $R(t)$, we then have (just as we did for wave functions),[43]

$$U(R(t)) = 1_{\mathcal{H}} - \frac{it\,\omega\cdot\mathbf{J}}{\hbar} + O(t^2) . \tag{6.20}$$

We can compare the group-theoretic commutator of two rotations with the composition taken both before and after applying the map to $\mathsf{U}(\mathcal{H})$; we have the equation

$$U(R(t))\,U\left(\tilde{R}(t)\right)\,U(R(t))^*\,U\left(\tilde{R}(t)\right)^* = U\left(R(t)\tilde{R}(t)R(t)^{-1}\tilde{R}(t)^{-1}\right) , \tag{6.21}$$

where on the left we have the commutator of elements of $\mathsf{U}(\mathcal{H})$, and on the right we have the image in $\mathsf{U}(\mathcal{H})$ of the commutator of elements of $\mathsf{SO}(3)$. Letting each rotation be infinitesimal of the same order, we get, by comparing terms at second order,

$$[\omega\cdot\mathbf{J}, \tilde{\omega}\cdot\mathbf{J}] = i\hbar(\omega_{(1)}\wedge\omega_{(2)})\cdot\mathbf{J} \tag{6.22}$$

---

[43] For the very discerning reader, the linear dependence on ω in (6.20) requires some explanation.

which is exactly analogous to (6.18) with **L** replaced by **J**. In components, this is

$$[J_i, J_j] = i\hbar \sum_k \varepsilon_{ijk} J_k \ . \tag{6.23}$$

This is an important result; whenever we have a representation of the rotation group on a Hilbert space, we get a trio of self-adjoint angular momentum operators, $\{J_i\}$, that obey the commutation relations (6.23) and generate the action of more general rotations via exponentiation in the sense of Stone's theorem. This is an instance of a fundamental relationship between representations of Lie groups and representations of Lie algebras.

6.4 Angular momentum multiplets

We have (at least partially) reduced problem of studying of rotations in quantum systems to the study of *representations of the angular momentum operators*:

Definition 6.4.1. A *representation of the angular momentum operators* is a Hilbert space, $\mathcal{H}$, equipped with an action of three self-adjoint operators $J_i : \mathcal{H} \to \mathcal{H}$, $i = 1, 2, 3$, satisfying the commutation relations (6.23).

Remark 6.4.2. This is equivalent to a representation of the Lie algebra $\mathfrak{so}(3)$ on $\mathcal{H}$. The difference is in the factor of $\hbar$ on the right hand side of (6.23), which can be removed by an appropriate rescaling of the J_i. Also, in some cases it is conventional for a representation of $\mathfrak{so}(3)$ to be described in terms of anti-self adjoint operators (operators obeying $A^* = -A$), in which case a factor of i is incorporated into the rescaling as well.

Definition 6.4.3. An *irreducible representation* of the angular momentum operators is a representation of the angular momentum operators for which there is no a proper subspace $\mathcal{H}_{\text{sub}} \subset \mathcal{H}$ with $J_i : \mathcal{H}_{\text{sub}} \to \mathcal{H}_{\text{sub}}$, *i.e.*, $\mathcal{H}$ contains no proper sub-representation of the angular momentum operators.

In **A11 Quantum Theory**, in the context of discussing orbital angular momentum for three-dimensional wave functions, you identified the structure of general irreducible representations of the angular momentum operators. Here we will recall the story in the general case. We define the *total angular momentum operator* $\mathbf{J}^2 = \mathbf{J} \cdot \mathbf{J}$. A short calculation shows that

$$[\mathbf{J}^2, J_i] = 0 \ , \tag{6.24}$$

so the action of the J_i operators preserves eigenspaces of $\mathbf{J}^2$. Since $\mathbf{J}^2$ is self-adjoint, we can choose to work in a basis of its eigenstates for any representation of the angular momentum operators, and so if $\mathcal{H}$ is an irreducible representation, then $\mathbf{J}^2$ must just act by a multiple of the identity on $\mathcal{H}$. We can give a completely explicit description of all finite-dimensional, irreducible representations if we furthermore choose to diagonalise J_3.

Theorem 6.4.4. The irreducible representations of the angular momentum operators are labeled by a non-negative half-integer $j = 0, \frac{1}{2}, 1, \ldots \in \frac{1}{2}\mathbb{N}$ known as the *spin* of the representation. Denote the Hilbert space admitting such a representation by $\mathcal{H}_{spin\, j}$. The dimension of $\mathcal{H}_{spin\, j}$ is $2j+1$ and $\mathbf{J}^2$ acts with eigenvalue $\hbar^2 j(j+1)$.

There is an orthonormal basis of $\mathcal{H}_{spin\, j}$ consisting of eigenvectors $|j,m\rangle$ of J_3 with $J_3 |j,m\rangle = \hbar m |j,m\rangle$ for $m = -j, -j+1, \ldots j-1, j$.

Proof. We introduce the ladder operators $J_\pm = J_1 \pm i J_2$, which commute with $\mathbf{J}^2$. We also can check that

$$[J_3, J_\pm] = \pm \hbar J_\pm \ . \tag{6.25}$$

This gives them the interpretation as raising and lowering operators for eigenvectors $|j,m\rangle$ of J_3 (with eigenvalue $\hbar m$, say):

$$J_3 \left(J_\pm |j,m\rangle \right) = \pm \hbar J_\pm |j,m\rangle + J_\pm J_3 |j,m\rangle = \hbar (m \pm 1) \left(J_\pm |j,m\rangle \right) \ . \tag{6.26}$$

Thus $J_\pm |j,m\rangle$ is a multiple of an eigenvector for J_3 with eigenvalue $\hbar(m \pm 1)$. The following then shows that the values for $|m|$ must be bounded.

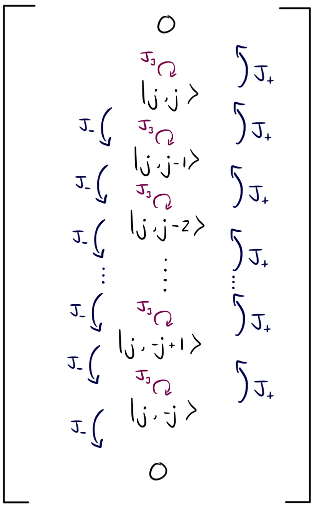


Figure 2. Depiction of irreducible representation of the angular momentum operators.

Lemma 6.4.5. Let $\mathbf{J}^2 |\psi\rangle = \lambda\hbar^2 |\psi\rangle$ and $J_3 |\psi\rangle = \hbar m |\psi\rangle$. Then for all $\phi \in \mathcal{H}$,

$$\langle J_\pm \phi | J_\pm \psi\rangle = \hbar^2 \left(\lambda - m(m \pm 1)\right) \langle \phi|\psi\rangle \quad \text{and} \quad \|J_\pm \psi\|^2 = \hbar^2 \left(\lambda - m(m\pm 1)\right) \|\psi\|^2 \,. \tag{6.27}$$

Proof. Observe from the angular momentum commutation relations that

$$J_+ J_- = \mathbf{J}^2 - J_3^2 + \hbar J_3 \,, \qquad J_- J_+ = \mathbf{J}^2 - J_3^2 - \hbar J_3 \,, \tag{6.28}$$

so the identities follow from

$$\langle J_- \phi | J_- \psi\rangle = \langle \phi | J_+ J_- \psi\rangle = \left\langle \phi \middle| (\mathbf{J}^2 - J_3^2 + \hbar J_3)\psi \right\rangle \,, \tag{6.29}$$

and using the eigenvalue relations (and similarly for the J_+ plus version of (6.29)). ■

Given that λ is fixed on an irreducible representation, $|m|$ cannot be too large as otherwise the norm squared of these states would be negative. The only way to avoid $|m|$ becoming arbitrarily large in the negative direction is if for some smallest value m_-, $J_- |\psi_{m_-}\rangle = 0$ where $J_3 |\psi_{m_-}\rangle = \hbar m_- |\psi_{m_-}\rangle$, which requires $\lambda = m_-(m_- - 1)$. The only way that $|m|$ can avoid becoming arbitrarily large in the positive direction is if analogously for some largest value m_+, $J_+|\psi_{m_+}\rangle = 0$, so $\lambda = m_+(m_+ + 1)$. To realise both situations at once, we need

$$\lambda = j(j+1) \,, \qquad m_- = -j \,, \qquad m_+ = j \,. \tag{6.30}$$

By construction $m_+ - m_- = 2j$ must be an integer (since starting with the $|\psi_{m_-}\rangle$ and acting repeatedly with J_+ we must arrive eventually at $|\psi_{m_+}\rangle$). Hence the constraints on the eigenvalues are as stated in the theorem.

To finish off the proof, we require that the J_3 eigenvalues be nondegenerate. This follows from irreducibility. Suppose that there are two linearly independent eigenvectors $|j, m; 1\rangle$ and $|j, m; 2\rangle$ that, without loss of generality, can be taken to be mutually orthogonal. Then it follows from the expressions above that $J_\pm^n |j, m; 1\rangle$ and $J_\pm^n |j, m; 2\rangle$ are orthogonal. Thus there will be two nontrivial J_i-invariant subspaces spanned by $J_\pm^n |j, m; 1\rangle$ and by $J_\pm^n |j, m; 2\rangle$, contradicting irreducibility. ■

We conclude with a few additional comments:

- If we are working in a definite irreducible representation of spin j, we might sometimes simply denote the state kets $|m\rangle$ to encode the J_3 eigenvalue.

- The basis $|j,m\rangle$ of $\mathcal{H}_{spin\, j}$ is unique up to an *overall normalisation* for the entire representation if we impose the normalisation conditions
$$|j,m\pm 1\rangle = \frac{J_\pm\,|j,m\rangle}{\hbar\sqrt{j(j+1)-m(m\pm 1)}}\,. \tag{6.31}$$
This definition ensures in particular that the states $|j,m\rangle$ all have the same norm, so if we choose a particular state, say $|j,j\rangle$, to be unit normalised and construct the rest of the representation by the action of J_-, then all of these states will be unit normalised.

- Important examples of representations with integer spin were given in **A11 Quantum Theory** in terms of spherical harmonics. These are angular momentum representations realised using the orbital angular momentum operators $\mathbf{L}$, which can be written in spherical polar coordinates as
$$L_\pm = i\hbar\, e^{\pm i\varphi}\left(\cot\theta\frac{\partial}{\partial\varphi}\pm i\frac{\partial}{\partial\theta}\right)\,, \qquad L_3 = -i\hbar\frac{\partial}{\partial\varphi}\,. \tag{6.32}$$
The total spin j is usually denoted by ℓ in this context and is required to be an integer. The wave functions $\Psi_\ell^m(\varphi,\theta)$ corresponding to the basis states $|\ell,m\rangle$ take the form
$$Y_\ell^m(\varphi,\theta) = P_\ell^m(\cos\theta)e^{im\varphi}\,, \tag{6.33}$$
where $P_\ell^m(x)$ are associated Legendre functions. The requirement that ℓ and m be integral follows from the need for $e^{im\varphi}$ to be single valued.

Example 6.4.6 (Spin 1/2). We saw that while half-integral spin is acceptable in the context of representations of the angular momentum operators, it doesn't arise in the context of orbital angular for three-dimensional wave functions. Let us investigate the simplest case: spin $j = 1/2$.
The discussion above gives an explicit realisation of this representation,
$$\mathcal{H}_{\text{spin}\,\frac{1}{2}} \cong \mathbb{C}^2 = \text{Span}\{\,|\tfrac{1}{2},\tfrac{1}{2}\rangle\,,|\tfrac{1}{2},-\tfrac{1}{2}\rangle\,\}\,. \tag{6.34}$$
Of course this is just our old friend the qubit. The above action of $J_\pm$ and hence J_1 and J_2 is determined by (6.31) for which in this case the denominator is just $\hbar$, and the eigenvalue condition determines J_3. It follows that in this basis we have
$$\mathbf{J} = \frac{\hbar}{2}\sigma\,, \tag{6.35}$$
where $\sigma = (\sigma_1,\sigma_2,\sigma_3)$ are the same Pauli spin matrices we met in our qubit discussion. Now let us consider a general rotation by some angle θ about an axis designated by the unit vector $\mathbf{n}$; we denote this by $R_{\mathbf{n}}(\theta)$. By Stone's theorem, this should be realised on our two-dimensional Hilbert space by the unitary matrix
$$U(R_{\mathbf{n}}(\theta)) =: U_{\mathbf{n}}(\theta) = \exp\left(-\frac{i\theta}{\hbar}\mathbf{n}\cdot\mathbf{J}\right) = \exp\left(-\frac{i\theta}{2}\mathbf{n}\cdot\sigma\right)\,. \tag{6.36}$$
An explicit computation of this matrix exponential yields a simple expression for the matrix that should represent the rotation,
$$U_{\mathbf{n}}(\theta) = \cos\left(\frac{\theta}{2}\right)1_{2\times 2} - i\sin\left(\frac{\theta}{2}\right)\mathbf{n}\cdot\sigma\,. \tag{6.37}$$
It is easy to confirm that these are unitary matrices, and in addition they are manifestly traceless, so are elements of $\mathrm{SU}(2)$. Indeed, by letting $\mathbf{n}$ range over the unit sphere in three dimensions and letting θ run from 0 to 2π, this gives a parameterisation of the most general element of $\mathrm{SU}(2)$. However, compared to rotations this is double counting! Rotating by θ around the axis defined by $\mathbf{n}$ is the same as rotating by $2\pi-\theta$ around the axis defined by $-\mathbf{n}$.
Indeed, for fixed $\mathbf{n}$, we see that setting $\theta = 2\pi$ doesn't give us back the identity, but rather minus the identity. It is only upon taking $\theta = 4\pi$ that our unitary matrix returns to the identity. So there is a two-to-one correspondence between the elements of $\mathrm{SU}(2)$ and the inequivalent rotations, *i.e.*, the elements of $\mathrm{SO}(3)$.

We are encountering in this example precisely a situation where our symmetry group (SO(3)) is implemented via a projective unitary representation that is not strictly a unitary representation of the group we started with. We can see this in terms of the group law. Consider the rotation $R_{\mathbf{n}}(\pi)$ that performs a half rotation about the axis **n**. Then performing this twice we have have

$$U(R_{\mathbf{n}}(\pi))\,U(R_{\mathbf{n}}(\pi)) = U_{\mathbf{n}}(2\pi) = -\mathbb{1}_{2\times 2}\ . \tag{6.38}$$

whereas if we compose the rotations before taking the map to unitary matrices, we have

$$U(R_{\mathbf{n}}(\pi)R_{\mathbf{n}}(\pi)) = U(R_{\mathbf{n}}(2\pi)) = U(R_{\mathbf{n}}(0)) = \mathbb{1}_{2\times 2}\ . \tag{6.39}$$

The sign difference is precisely the type of “extra phase” that is allowed for projective representations!
It turns out that this example is indicative of the general story for half-integer-spin representations. These are projective unitary representations of SO(3) that do not lift to unitary representations of SO(3). Rather, they correspond to unitary representations of SU(2), where the relation between the two groups is by a quotient,

$$\mathbb{P}\,\mathrm{SU}(2) := \mathrm{SU}(2)/\{\pm 1\} \cong \mathrm{SO}(3)\ . \tag{6.40}$$

In the case of orbital angular momentum, there is *manifestly* a representation of the honest rotation group via the action on wave functions; consequently only integer spin can occur.

Remark 6.4.7. There is a beautiful observation to make here that I cannot help but include for your entertainment (I hope). As was observed above, we have a realisation of SU(2) by a choice of unit vector in $\mathbb{R}^3$ and an angle $\theta \in 2\pi$. This gives us a realisation of SU(2) as a *circle fibration over the two-sphere* (you can imagine a circle corresponding to the choice of angle sitting over each point on the two-sphere corresponding to the choice of unit vector). This is what’s known as the *Hopf fibration*, which realises the three sphere $S^3 \cong \mathrm{SU}(2)$ as a circle fibration over S^2. The rotation group $SO(3)$ then gets identified as the quotient space $S^3/\mathbb{Z}_2$, with $\mathbb{Z}_2$ acting as the antipodal map.

Chapter 7

Intrinsic Spin and Addition of Angular Momentum

When we model a quantum mechanical particle or system that has some internal structure, the action of the rotation group can be more complicated than what we have in the case of wave functions. We could describe such a system in terms of its center of mass, which will be labelled by a point in $\mathbb{R}^3$, as well as some internal structure whose quantum mechanical configurations are encoded in a Hilbert space $\mathcal{H}_{\text{internal}}$. The full Hilbert space for such a structured particle will then given by

$$\mathcal{H} = L^2(\mathbb{R}^3) \otimes \mathcal{H}_{\text{internal}} \,. \tag{7.1}$$

We will then have a (projective) unitary representation of $\mathrm{SO}(3)$ on this Hilbert space generated by some *total angular momentum operators* $\mathbf{J}$. At the infinitesimal level, we know that $\mathbf{J}$ will act on the center of mass wave function as the orbital angular momentum operator $\mathbf{L}$, while we can introduce operators $\mathbf{S}$ that describe the action on $\mathcal{H}_{\text{internal}}$,

$$\mathbf{J} = \mathbf{L} \otimes 1_{\mathcal{H}_{\text{internal}}} + 1_{L^2(\mathbb{R})} \otimes \mathbf{S} \,, \tag{7.2}$$

where the $\mathbf{L}$ and $\mathbf{S}$ operators commute and each satisfy the angular momentum commutation relations,

$$[L_i, L_j] = i\hbar \sum_k \varepsilon_{ijk} L_k \,, \qquad [S_i, S_j] = i\hbar \sum_k \varepsilon_{ijk} S_k \,. \tag{7.3}$$

If we are describing *elementary* particles, we expect $\mathcal{H}_{\text{internal}}$ to support an irreducible angular momentum representation; otherwise we would think of the different subrepresentations as corresponding to (detectably!) different versions of the elementary particle, and therefore we would call them different types of elementary particles. If a particle has $\mathcal{H}_{\text{spin}\, j}$ as its internal Hilbert space with the attendant operators $\mathbf{S}$ describing rotations, then we say the particle itself has *intrinsic spin j* (though often we simply shorten this to say the particle has spin j). This intrinsic spin is the quantity that appears in the spin statistics theorem of Chapter 4.

7.1 Addition of angular momentum

To understand the full action of rotations on a system with both orbital angular momentum and intrinsic spin, we will have to understand the action of the total angular momentum operators in a system where we initially understand the action of the components $\mathbf{L}$ and $\mathbf{S}$ separately. Similarly, if we have a system of n particles each of which has some angular momentum operator $\mathbf{J}_i$ acting on its single-particle Hilbert space, then we will want to understand the total angular momentum[44] $\mathbf{J} = \mathbf{J}_{(1)} + \ldots + \mathbf{J}_{(n)}$ acting on the composite Hilbert space starting with an understanding of the individual angular momentum representations. This procedure is, in the physics literature, usually referred to as the *addition of angular momentum*. We first pose the problem in the context of there being two constituent representations of angular momentum.

Problem 7.1.1. *Consider the Hilbert space $\mathcal{H} \cong \mathcal{H}_1 \otimes \mathcal{H}_2$ where $\mathcal{H}_i$ supports an irreducible representation of the angular momentum operators $\mathbf{J}^{(i)}$ with spin j_i. We have $\dim \mathcal{H} = (2j_1+1)(2j_2+1)$. How does this composite system decompose into irreducible representations of the total angular momentum operator $\mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)}$?*

The answer is given in the following Proposition.

Proposition 7.1.2. Under the action of the total angular momentum, the tensor product $\mathcal{H} \cong \mathcal{H}_1 \otimes \mathcal{H}_2$ of irreducible representations with spins j_1 and j_2, respectively, decomposes into irreducible representations according to

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \bigoplus_{J=|j_1-j_2|}^{j_1+j_2} \mathcal{H}_{\text{spin}\, J} \,. \tag{7.4}$$

---

[44] Here and in what follows we drop the explicit tensor notation for these sums of operators acting on different tensor factors of a composite Hilbert space.

Thus we can find an orthonormal basis for this tensor product that we denote by $|J,M\rangle$ with $J = |j_1 - j_2|, |j_1 - j_2| + 1, \ldots, j_1 + j_2 - 1, j_1 + j_2$ and $M = -J, -J+1, \ldots, J-1, J$ which obey[45]

$$(\mathbf{J}^{(1)})^2\,|J,M\rangle \;=\hbar^2 j_1(j_1+1)\,|J,M\rangle\;, \qquad \mathbf{J}^2\,|J,M\rangle \;=\hbar^2 J(J+1)\,|J,M\rangle\;, \tag{7.5}$$

$$(\mathbf{J}^{(2)})^2\,|J,M\rangle \;=\hbar^2 j_2(j_2+1)\,|J,M\rangle\;, \qquad J_3\,|J,M\rangle \;=\hbar M\,|J,M\rangle\;. \tag{7.6}$$

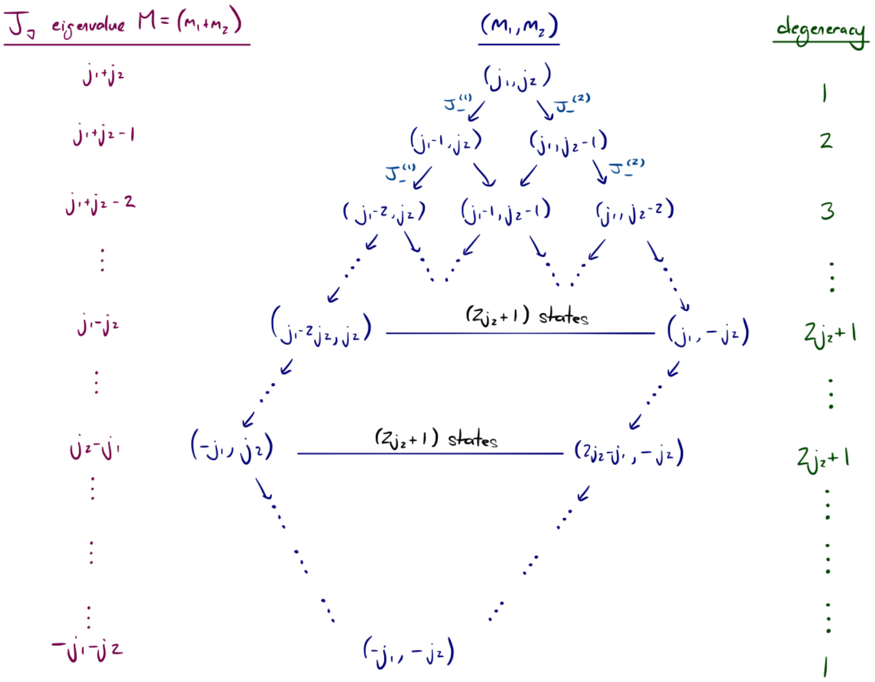


Figure 3. Depiction of states in the tensor product of irreducible representations of angular momentum with spins j_1 and j_2, where we assume $j_1 \geqslant j_2$.

Proof. The situation described in the following proof is illustrated in Figure 3, which hopefully makes the combinatorics of the proof easier to follows.

We assume, without loss of generality, that $j_1 \geqslant j_2$. We then proceed inductively, first finding the representation of highest total spin $\mathcal{H}_{\text{spin }(j_1+j_2)}$ inside $\mathcal{H}$, then in the complement $\mathcal{H}^{\perp}_{\text{spin }(j_1+j_2)} \subset \mathcal{H}$ finding the representation of next highest total spin, and so on.

We will write our pure-tensor basis vectors as $|j_1,m_1\rangle \otimes |j_2,m_2\rangle = |m_1;m_2\rangle$,[46] which as a reminder, obey

$$\begin{aligned}(\mathbf{J}^{(1)})^2\,|m_1;m_2\rangle &= \hbar^2 j_1(j_1+1)\,|m_1;m_2\rangle\;, \qquad & J_3^{(1)}\,|m_1;m_2\rangle &= \hbar\,m_1\,|m_1;m_2\rangle\;,\\ (\mathbf{J}^{(2)})^2\,|m_1;m_2\rangle &= \hbar^2 j_2(j_2+1)\,|m_1;m_2\rangle\;, \qquad & J_3^{(2)}\,|m_1;m_2\rangle &= \hbar\,m_2\,|m_1;m_2\rangle\;.\end{aligned} \tag{7.7}$$

These are evidently already eigenvectors of $J_3 = (J^{(1)})_3 + (J^{(2)})_3$, with eigenvalues $M = m_1 + m_2$. There is a unique state with maximum $M = j_1 + j_2$, which we must be able to identify with the "top" state in a spin $j_1 + j_2$ representation; we can therefore identify

$$|J = j_1 + j_2, M = j_1 + j_2\rangle \;=\; |j_1;j_2\rangle\;. \tag{7.8}$$

---

[45] These states are often written as $|j_1,j_2;J,M\rangle$ to make manifest the constituent spins that are being combined. We will leave these implicit to avoid overly burdensome notation whenever possible.

[46] As with the previous basis states, these will sometimes be labelled $|j_1,m_1;j_2,m_2\rangle$ to indicate the constituent spins. We will avoid this when possible to minimise notational clutter.

The rest of the spin $J = j_1+j_2$ representation can be recovered by acting with the total lowering operator $J_- = J_-^{(1)}+J_-^{(2)}$. Normalising these using (6.31), we obtain the states $|J = j_1 + j_2, M\rangle$ for $M < J$. These necessarily constitute a complete spin $J = j_1 + j_2$ subrepresentation in $\mathcal{H}$.

Next we consider the states with $M = j_1+j_2-1$. There are two linearly independent such states, $|j_1 - 1; j_2\rangle$ and $|j_1; j_2 - 1\rangle$. One combination of these occurs in the spin j_1+j_2 representation defined above as $|J = j_1 + j_2, M = j_1 + j_2 - 1\rangle$. Taking a vector in the orthogonal complement, we get a state vector that is necessarily the top state in a spin j_1+j_2-1 representation, since acting with a raising operator must give zero, else this would be part of the previous spin $j_1 + j_2$ representation. We therefore denote a normalised element of this orthogonal complement by $|J = j_1 + j_2 - 1, M = j_1 + j_2 - 1\rangle$. Again acting with $J_-^{(1)} + J_-^{(2)}$ on this state generates a full irreducible subrepresentation, this time of spin $j_1 + j_2 - 1$.

If $j_2 = \frac{1}{2}$, then the degeneracy for M eigenvalues with $j_1 + j_2 - 1 \geqslant M \geqslant -(j_1 + j_2 - 1)$ is just two, corresponding to the $m_2 = \pm\frac{1}{2}$ states, and so must be spanned by the corresponding M-eigenstates of the two multiplets we have just found. In this case we would be done.

If $j_2 > \frac{1}{2}$, then the degeneracy for the $M = j_1 + j_2 - 2$ eigenvalue is three with $(m_1, m_2) = (j_1, j_2 - 2), (j_1 - 1, j_2 - 1)$ or $(j_1 - 2, j_2)$. Thus, as before, there is a nontrivial orthogonal normalised vector $|j_1 + j_2 - 2, j_1 + j_2 - 2\rangle$ orthogonal to those $M = j_1 + j_2 - 2$ eigenvalues of total spin $j_1 + j_2$ and $j_1 + j_2 - 1$, unique up to a phase. This gives rise to a spin $J = j_1 + j_2 - 2$ representation by lowering.

In general, the degeneracy states with J_3 eigenvalue M is given by $1 + j_1 + j_2 - |M|$ for $|M| \geqslant j_1 - j_2$, but is $2j_2 + 1$ otherwise as it cannot exceed the number of choices $2j_2 + 1$ for m_2 (see Figure 3). So we can carry on by induction, generating a new multiplet at each stage, until we eventually produce all the angular momentum multiplets with spins from $J = j_1 - j_2$ to $J = j_1 + j_2$, as required. This gives a total of $2j_2 + 1$ irreducible representations ($2j_2 + 1$ being the maximal degeneracy of the M eigenvalue, realised for $|M| \leqslant j_1 - j_2$). ■

Before moving on, we will look at the simplest case of addition of angular momentum: the tensor product of two spin 1/2 systems.

Example 7.1.3 (Two qubits). We can be very explicit in examining the two-qubit system,

$$\mathcal{H} = \mathcal{H}_{\text{spin } \frac{1}{2}}^{\otimes 2} . \tag{7.9}$$

If we adopt the basis for the spin-$\frac{1}{2}$ qubit from the last chapter, we have as our basis for the tensor product Hilbert space $\left|\pm\frac{1}{2}; \pm\frac{1}{2}\right\rangle$, with the signs chosen independently. To make things easier on the eyes, we will adopt the notation $|\pm\pm\rangle$. If we organise these by J_3 eigenvalue $M = m_1 + m_2$, we have

$$\begin{array}{lccc} M = +1: & & |{+}{+}\rangle & \\ M = 0: & |{+}{-}\rangle & & |{-}{+}\rangle \\ M = -1: & & |{-}{-}\rangle & \end{array} \tag{7.10}$$

Starting with the top state, we produce the following states in the spin-one subrepresentation of this system,

$$|1, 1\rangle = |{+}{+}\rangle \ , \quad |1, 0\rangle = \frac{|{+}{-}\rangle + |{-}{+}\rangle}{\sqrt{2}} \ , \quad |1, -1\rangle = |{-}{-}\rangle \ . \tag{7.11}$$

There is an additional spin-zero state (so a rotationally-invariant state),[47]

$$|0, 0\rangle = \frac{|{+}{-}\rangle - |{-}{+}\rangle}{\sqrt{2}} . \tag{7.12}$$

It may be worth remarking that the spin-one representation consists of bosonic (symmetric) states, while the spin-zero representation is the one fermionic (anti-symmetric) state in this tensor product. Indeed, the total angular momentum operators can be seen (by inspection) to commute with the action of permutations on the n-fold tensor product of identical representations of angular momentum, which means that the bosonic and fermionic subspaces will always transform among themselves under rotations.

Example 7.1.4 (Everything from spin one half)**.** The previous example suggests a general construction of the spins j representation for any j using spin-$\frac{1}{2}$ representations as building blocks. Consider the n-fold symmetric tensor product of the qubit Hilbert space $\odot^n\mathcal{H}_{\text{spin }\frac{1}{2}}$. By our results in Chapter 4, this has dimension

$$\dim\left(\odot^n\mathcal{H}_{\text{spin }\frac{1}{2}}\right) = n+1\,. \tag{7.13}$$

If we consider the state $|++\cdots+\rangle$, this has

$$J_3^{(tot)}\,|++\cdots+\rangle = \frac{\hbar n}{2}\,|++\cdots+\rangle\ , \tag{7.14}$$

so this must be an element of a representation with spin greater than or equal to $n/2$. But on dimensionality grounds, this can only be a representation of spin $n/2$ exactly, and indeed it is clear that this state is the top state of its angular momentum representation. Thus, if you like, you can think of any irreducible representation of angular momentum in terms of an appropriate number of identical (bosonic) qubits. This can prove a useful mental heuristic for these representations.

7.2 Clebsch–Gordan coefficients

Equation (7.4) tells us—in general terms—how the tensor product of irreducible angular momentum representations will transform under the total angular momentum. However, to work with these composite systems and do calculations in practice, one needs to be able to concretely construct and manipulate the elements of the different irreducible representations appearing in the direct sum on the right hand side of that equation. In practice, this usually means having expressions for the precise linear combinations of the states $|m_1;m_2\rangle$ that constitute the states $|J,M\rangle$ and *vice versa*. To this end we make the following definition.

Definition 7.2.1. The *Clebsch–Gordan coefficients* $C_{j_1j_2}(J,M;m_1,m_2)$ are defined by

$$|J,M\rangle = \sum_{\substack{m_1,m_2\\ m_1+m_2=M}} C_{j_1j_2}(J,M;m_1,m_2)\,|m_1;m_2\rangle\ , \tag{7.15}$$

so these are the coefficients of the expansion of our orthonormal basis of states in the (J,M) basis in terms of those in the separable, (m_1,m_2) basis.

The coefficients defined as such are not uniquely specified defined, because there is some choice of overall phases in the states for each irreducible representation involved. However, this freedom can be fixed by requiring that the Clebsch–Gordan coefficients be real along with an additional convention that we will describe momentarily.

Because both bases are orthonormal, we can deduce a number of useful expressions involving the Clebsch–Gordan coefficients. For example, we can realise them explicitly in terms of inner products of the form

$$C_{j_1j_2}(J,M;m_1,m_2) = \langle m_1;m_2|J,M\rangle\ . \tag{7.16}$$

If we adopt the conventions mentioned above to ensure reality of the Clebsch–Gordan coefficients, then we will also have

$$C_{j_1j_2}(J,M;m_1,m_2) = \langle J,M|m_1;m_2\rangle\ . \tag{7.17}$$

---

[47] You might recognise this as the EPR state from our brief discussion of entanglement.

The completeness relation for our Hilbert space leads to the following identity,

$$\begin{aligned} 1 &= \langle J, M | J, M \rangle \ , \\ &= \sum_{m_1, m_2} \langle J, M | m_1; m_2 \rangle \langle m_1; m_2 | J, M \rangle \ , \\ &= \sum_{m_1, m_2} |C_{j_1 j_2}(J, M; m_1, m_2)|^2 \ . \end{aligned} \tag{7.18}$$

Analogously, the completeness relation for the $|J, M\rangle$ states gives

$$\begin{aligned} 1 &= \langle m_1; m_2 | m_1; m_2 \rangle \ , \\ &= \sum_{J,M} \langle m_1, m_2 | J, M \rangle \langle J, M | m_1; m_2 \rangle \ , \\ &= \sum_{J,M} |C_{j_1 j_2}(J, M; m_1, m_2)|^2 \ . \end{aligned} \tag{7.19}$$

More generally, we have

$$\begin{aligned} \delta_{JJ'} \delta_{MM'} &= \sum_{m_1, m_2} \langle J, M | m_1; m_2 \rangle \langle m_1; m_2 | J', M' \rangle \ , \\ &= \sum_{m_1, m_2} \overline{C_{j_1 j_2}(J, M; m_1, m_2)} C_{j_1 j_2}(J', M'; m_1, m_2) \ , \\ \delta_{m_1 m_1'} \delta_{m_2 m_2'} &= \sum_{J,M} \langle m_1, m_2 | J, M \rangle \langle J, M | m_1'; m_2' \rangle \ , \\ &= \sum_{J,M} C_{j_1 j_2}(J, M; m_1, m_2) \overline{C_{j_1 j_2}(J, M; m_1', m_2')} \ . \end{aligned} \tag{7.20}$$

We will now see take a look at how the computation of these coefficients works out in some simple examples.

Example 7.2.2. Let $j_2 = \frac{1}{2}$ with $j_1 \neq 0$ an arbitrary spin. Then the highest-spin state takes the usual form $|j_1 + \frac{1}{2}, j_1 + \frac{1}{2}\rangle = |j_1; \frac{1}{2}\rangle$. Acting with the total lowering operators, we get

$$\begin{aligned} |j_1 + \tfrac{1}{2}, j_1 - \tfrac{1}{2}\rangle &= \frac{1}{\hbar\sqrt{2j_1 + 1}} J_- \, |j_1 + \tfrac{1}{2}, j_1 + \tfrac{1}{2}\rangle \ , \\ &= \frac{1}{\hbar\sqrt{2j_1 + 1}} \left(J_-^{(1)} + J_-^{(2)} \right) |j_1; \tfrac{1}{2}\rangle \ , \\ &= \frac{1}{\sqrt{2j_1 + 1}} \left(\sqrt{2j_1} \, |j_1 - 1; \tfrac{1}{2}\rangle + |j_1; -\tfrac{1}{2}\rangle \right) \ . \end{aligned} \tag{7.21}$$

We can identify the most general (normalised) orthogonal complement in the $M = j_1 - \frac{1}{2}$ eigenspace as

$$|j_1 - \tfrac{1}{2}, j_1 - \tfrac{1}{2}\rangle \sim \frac{1}{\sqrt{2j_1 + 1}} \left(|j_1 - 1; \tfrac{1}{2}\rangle - \sqrt{2j_1} \, |j_1; -\tfrac{1}{2}\rangle \right) \ . \tag{7.22}$$

At this point we can use some of our freedom in introducing phases to fix the overall phase of this state. Demanding that the Clebsch–Gordan coefficients (so the coefficients of the expansion) be real gives us

$$|j_1 - \tfrac{1}{2}, j_1 - \tfrac{1}{2}\rangle = \pm \frac{1}{\sqrt{2j_1 + 1}} \left(\sqrt{2j_1} \, |j_1; -\tfrac{1}{2}\rangle - |j_1 - 1; \tfrac{1}{2}\rangle \right) \ . \tag{7.23}$$

To fix the final sign ambiguity, one may adopt a standard convention known as the *Condon–Shortley convention*. This amounts to declaring that

$$C_{j_1 j_2}(J, J; j_1, J - j_1) > 0 \ . \tag{7.24}$$

In the above, this selects the plus sign, and with that highest state fixed, we can produce the rest of the spin $j_1 - \frac{1}{2}$ multiplet by acting with J_-. (Note that this convention depends on the order of the two constituent spins j_1 and j_2.)

One can read off the Clebsch–Gordan coefficients from the resulting expressions for our states, *e.g.*, from (7.21) and (7.23). We have:

$$\begin{aligned}
C_{j_1\,\frac{1}{2}}\left(j_1+\tfrac{1}{2}, j_1+\tfrac{1}{2}; j_1, \tfrac{1}{2}\right) &= 1\,,\\
C_{j_1\,\frac{1}{2}}\left(j_1+\tfrac{1}{2}, j_1-\tfrac{1}{2}; j_1, -\tfrac{1}{2}\right) &= \frac{1}{\sqrt{2j_1+1}}\,,\\
C_{j_1\,\frac{1}{2}}\left(j_1+\tfrac{1}{2}, j_1-\tfrac{1}{2}; j_1-1, \tfrac{1}{2}\right) &= \sqrt{\tfrac{2j_1}{2j_1+1}}\,,\\
C_{j_1\,\frac{1}{2}}\left(j_1-\tfrac{1}{2}, j_1-\tfrac{1}{2}; j_1, -\tfrac{1}{2}\right) &= \sqrt{\tfrac{2j_1}{2j_1+1}}\,,\\
C_{j_1\,\frac{1}{2}}\left(j_1-\tfrac{1}{2}, j_1-\tfrac{1}{2}; j_1-1, \tfrac{1}{2}\right) &= \frac{-1}{\sqrt{2j_1+1}}\,.
\end{aligned} \tag{7.25}$$

You can look up tables of Clebsch–Gordan coefficients in textbooks and online, but you should learn to love them and practice deriving some!

Example 7.2.3 (Application to hydrogen energy levels). An important application of this general story arises in the analysis of atoms. Let us start with a single-electron atom (a.k.a., a Hydrogen-like atom). We model this as an electron moving in an external Coulomb potential, and you have studied the corresponding stationary state wave functions in **A11 Quantum Theory**. However, the electron has intrinsic spin $1/2$, so we should really think of its Hilbert space as being the tensor product

$$\mathcal{H}_{\text{electron}} \cong L^2(\mathbb{R}^3) \otimes \mathcal{H}_{\text{spin}\,1/2}\,. \tag{7.26}$$

A general state in the electron Hilbert space will then consist of *two wavefunctions*, one for each of the possible internal spin states

$$|\psi\rangle = \left|\psi_+ \otimes +\tfrac{1}{2}\right\rangle + \left|\psi_- \otimes -\tfrac{1}{2}\right\rangle\,, \tag{7.27}$$

where $\left|\pm\frac{1}{2}\right\rangle$ are the $m_s = \pm\frac{1}{2}$ intrinsic spin eigenstates for, say, the S_3 operator, and $\psi_\pm(\mathbf{x})$ give the wave functions for the situation that the electron has $m_s = \pm\frac{1}{2}$. In non-relativistic quantum mechanics, the Schrödinger equation does not mix the two m-values, so the component wave functions $\psi_\pm(\mathbf{x})$ must both individually satisfy the same Schrodinger equation.[48]

If we take the wavefunctions $\psi_\pm(\mathbf{x})$ to be the stationary state wave functions you met for the Hydrogen atom previously, we end up with stationary states that we write as

$$|n, \ell, m_\ell; m_s\rangle = f_{n\ell}(r) Y_\ell^{m_\ell}(\theta, \varphi) \otimes |m_s\rangle\,, \tag{7.28}$$

where on the right we have separated out the explicit wave function part in front and left only the intrinsic spin state in the ket. Here, as elsewhere, the $Y_\ell^{m_\ell}(\theta, \varphi)$ are the spherical harmonics with eigenvalues $\hbar^2\ell(\ell+1)$ for $\mathbf{L}^2$ and $\hbar m_\ell$ for L_3, and $f_{n\ell}(r)$ is a (particular) polynomial in r of degree $n-1$ multiplied by $\exp(-Zr/na_0)$. These states simultaneously diagonalise $\mathbf{L}^2$, $\mathbf{S}^2$, L_3, S_3, and H, with energy

$$E_n = \frac{E_0}{n^2}\,, \quad E_0 = -\frac{Z^2 q_e^2}{2a_0}\,. \tag{7.29}$$

The energy eigenstates depend only on the *principal quantum number*, n, and for a given n there are states with $\ell = 0, \ldots, n-1$, and for each ℓ there are $2\ell+1$ different values of m_ℓ. Each such wave function also occurs twice, once for each of the two values of m_s. Adding everything up, there is a degeneracy of $2n^2$ for the energy level E_n.

Now for various purposes (we will see more on this in future chapters) it can prove useful to adopt a basis of states that diagonalises the total angular momentum operator $\mathbf{J}^2$ (where $\mathbf{J} = \mathbf{L} + \mathbf{S}$). Following our previous discussion of addition of angular momentum, we can find a basis of states that does this with eigenvalue $j(j+1)\hbar^2$, where $j = \ell \pm \frac{1}{2}$ (or $j = \frac{1}{2}$ if $\ell = 0$). These are given by the linear combinations,

$$\left|n,\ell,j,m_j\right\rangle = \sum_{\substack{m_\ell,m_s\\ m_\ell+m_s=m_j}} C_{\ell,\frac{1}{2}}(j,m_j;m_\ell,m_s)\left|n,\ell,m_\ell;m_s\right\rangle \ . \tag{7.30}$$

where $C_{\ell,\frac{1}{2}}(j,m_j;m_\ell,m_s)$ are of course the Clebsch–Gordan coefficients. This gives us a basis of stationary states that are eigenvectors for H, $\mathbf{L}^2$, $\mathbf{J}^2$, and J_3, the latter with eigenvalue m_j, but not for L_3 and S_3.
There is a standard nomenclature for the electron stationary states in this basis. The set of states with given values of n, ℓ, and j are known as $n\ell_j$ orbitals. For historical reasons, instead of writing the numerical value for $\ell = 0,1,2,3,4,\ldots$, one normally substitutes the letters $s,p,d,f,g,\ldots$ respectively (continuing alphabetically). Each such orbital contains $2j+1$ states with $m_j = -j,-j+1,\ldots,j-1,j$. So, for example:

- For all n, we have the $\ell = 0$ states (s orbitals) which appear only in the orbital $ns_{\frac{1}{2}}$ with the two states corresponding to $m_j = m_s = \pm\frac{1}{2}$.
- For $n \geqslant 2$ we can have $\ell = 1$ states (p orbitals) which arise in the configuration $np_{\frac{1}{2}}$ with two states or $np_{\frac{3}{2}}$ with four states, giving a total of 6 states for the np orbitals.
- In general, we have orbits $n\ell_{\ell\pm\frac{1}{2}}$ with $n \geqslant \ell+1$, and a similar counting for the $n\ell$ type of energy level gives a total of $2(\ell-\frac{1}{2})+1+2(\ell+\frac{1}{2})+1 = 4\ell+2$ states.

Remark 7.2.4. While we have reorganised the Hydrogen stationary states so as to diagonalise the total angular momentum operator, it may not yet be clear that this is a superior basis to choose. We will see in some examples in the next chapter that when we start considering *corrections* to the Hydrogen atom Hamiltonian arising from more subtle physical effects, these will often lead to a preference for one basis over another.

7.3 Irreducible tensor operators and the Wigner–Eckart theorem

Just as the Hilbert space of a rotational system can be organised in terms of the action of the angular momentum operators, so too can many *operators* in such systems. It then turns out that the matrix elements of operators with definite angular momentum properties (the so-called *irreducible tensor operators*, see below) are controlled by the same rules of addition of angular momentum that we have just developed.

We define an action of the rotation group on operators by conjugation by the corresponding unitaries,

$$A \xrightarrow{R} A' = U(R)AU(R)^* \ . \tag{7.31}$$

This definition is arranged so that under a combined action of rotations on states and operators, matrix elements remain fixed. (The idea is that if we simultaneously rotate the state of our system and the measuring apparatus (the operators), then the corresponding measurements/matrix elements should be invariant.)

$$\langle\psi|A|\varphi\rangle \xrightarrow{R} \langle U(R)\psi|A'|U(R)\varphi\rangle = \langle\psi|U(R)^*U(R)AU(R)^*U(R)|\varphi\rangle = \langle\psi|A|\varphi\rangle \ . \tag{7.32}$$

From this we can infer the transformation of an operator under an infinitesimal rotation, which is formulated in terms of the angular momentum operators,

$$U(R)AU(R)^* \approx (1 - \frac{i\varepsilon}{\hbar}\omega\cdot\mathbf{J})A(1+\frac{i\varepsilon}{\hbar}\omega\cdot\mathbf{J}) \approx A - \frac{i\varepsilon}{\hbar}[\omega\cdot\mathbf{J},A] \ . \tag{7.33}$$

---

[48] In a relativistic setting, the usual Schrödinger equation is replaced by the *Dirac equation*, which does mix up the different spin states. This is beyond the scope of our course.

Thus we identify the infinitesimal transformation $\delta_\omega A$ of an operator A with the commutator with the angular momentum operators,

$$\delta_\omega A = -\frac{i}{\hbar}[\omega \cdot \mathbf{J}, A] . \tag{7.34}$$

7.3.1 Vector operators

A number of operators that appear frequently in rotational systems have good reason to transform nicely under rotations. In particular, the position operators $\mathbf{X}$ or the momentum operators $\mathbf{P}$ should, in principle, *transform as vectors under rotations*. More precisely, they will transform according to,[49]

$$U(R)\mathbf{V}U(R)^* = R^{-1}\mathbf{V} , \qquad U(R)(\mathbf{n} \cdot \mathbf{V})U(R)^* = (R\mathbf{n}) \cdot \mathbf{V} , \tag{7.35}$$

Notice that with this definition, if we act with rotations on the *states* of a system while leaving the (vector) operators fixed,[50] then we have

$$\langle \psi | \mathbf{V} | \varphi \rangle \xrightarrow{R} \langle \psi | U(R)^* \mathbf{V} U(R) | \varphi \rangle = \big\langle \psi \big| U(R^{-1}) \mathbf{V} U(R^{-1})^* \big| \varphi \big\rangle = R \langle \psi | \mathbf{V} | \varphi \rangle , \tag{7.36}$$

so our measured expectation values/matrix elements transform by the rotation matrix R. Working infinitesimally, we have

$$\delta_\omega(\mathbf{n} \cdot \mathbf{V}) = (\omega \wedge \mathbf{n}) \cdot \mathbf{V} , \tag{7.37}$$

so in terms of angular momentum operators, we have

$$[\omega \cdot \mathbf{J}, \mathbf{n} \cdot \mathbf{V}] = i\hbar(\omega \wedge \mathbf{n}) \cdot \mathbf{V} . \tag{7.38}$$

we recognise this as the component-free expression for the commutation relations for the angular momentum operators, with some $\mathbf{J}$'s replaced by $\mathbf{V}$'s. Indeed working in components, we arrive at the following definition.

Definition 7.3.1. A *vector operator* is any triple of operators $\mathbf{V}$ in a rotational system that obey the commutation relations

$$[J_i, V_i] = i \sum_k \varepsilon_{ijk} V_k . \tag{7.39}$$

As an examples, we have (by direct computation) that the position, momentum, and angular momentum operators are all vector operators in this sense. More generally, operators can transform as irreducible representations (of arbitrary spin) of the rotation group/angular momentum operators. We make the following definition.

Definition 7.3.2. An *irreducible tensor operator operator of spin k* is a $(2k+1)$-tuple of operators $T_q^{(k)}$ for $q = -k, -k+1, \dots, k-1, k$ in a rotational system that obey the commutation relations

$$\begin{aligned} [J_3, T_q^{(k)}] &= \hbar\, q\, T_q^{(k)} , \\ [J_\pm, T_q^{(k)}] &= \hbar\sqrt{k(k+1) - q(q \pm 1)}\, T_{q\pm1}^{(k)} . \end{aligned} \tag{7.40}$$

Note that for the case $k = 1$, the basis $T_q^{(1)}$ is related to the Cartesian basis according to

$$T_0^{(1)} = V_3 , \qquad T_{\pm1}^{(1)} = \mp \frac{V_1 \pm iV_2}{\sqrt{2}} , \tag{7.41}$$

and with this change of basis the commutation relations in (7.39) and (7.40) coincide. The key result that motivates our definition of general tensor operators is the following theorem, which allows for a drastic simplification in the computation of matrix elements of tensor operators.

---

[49] The R^{-1} rather than R here is coming from the same place as the R^{-1} appearing in the action of rotations on wavefunctions.

[50] We think of this as rotating the state of our system while leaving the measuring apparatus/laboratory equipment fixed

Theorem 7.3.3 (Wigner–Eckart). The matrix elements of an irreducible tensor operator of spin k with respect to angular momentum eigenstates are given by

$$\left\langle \alpha; J, M \middle| T^{(k)}_q \middle| \beta; j, m \right\rangle = C_{j,k}(J, M; m, q) \left\langle \alpha; J \middle\| T^{(k)} \middle\| \beta; j \right\rangle \ . \tag{7.42}$$

Here α and β represent additional labels on the states that are not affected by the action of angular momentum operators. The double-bracketed object $\left\langle \alpha; J \middle\| T^{(k)} \middle\| \beta; j \right\rangle$ on the right hand side is called the *reduced matrix element*, and is some number that doesn't depend on the labels m, M, and q.

The punchline here is that by symmetry, the matrix elements of tensor operators between definite angular momentum states are determined entirely up to a single constant for each choice of the three angular momentum multiplets being coupled (two irreps as states and the choice of tensor operator), with the dependence on the particular states/elements of the angular momentum multiplets being entirely encoded in Clebsch–Gordan coefficients. In practice, the overall constant can then usually be evaulated by making a convenient choice of M, m, and q.

Proof. We consider the following matrix elements:

$$\left\langle \alpha; J, M \middle| [J_\pm, T^{(k)}_q] \middle| \beta; j, m \right\rangle = \hbar\sqrt{k(k+1) - q(q \pm 1)} \left\langle \alpha; J, M \middle| T^{(k)}_{q\pm 1} \middle| \beta; j, m \right\rangle \ , \tag{7.43}$$

which we can also rewrite by acting with the $J_\pm$ operators in the commutators to the left and to the right on the states, which yields

$$\hbar\sqrt{J(J+1) - M(M \mp 1)} \left\langle \alpha; J, M \mp 1 \middle| T^{(k)}_q \middle| \beta; j, m \right\rangle - \hbar\sqrt{j(j+1) - m(m \pm 1)} \left\langle \alpha; J, M \middle| T^{(k)}_q \middle| \beta; j, m \pm 1 \right\rangle \ . \tag{7.44}$$

Now observe that for the Clebsch–Gordan coefficients, we have

$$\begin{aligned} \langle J, M | J_\pm | j_1, m_1; j_2, m_2 \rangle &= \hbar\sqrt{j_1(j_1+1) - m_1(m_1 \pm 1)} \, \langle J, M | j_1, m_1 \pm 1; j_2, m_2 \rangle \\ &\quad + \hbar\sqrt{j_2(j_2+1) - m_2(m_2 \pm 1)} \, \langle J, M | j_1, m_1; j_2, m_2 \pm 1 \rangle \\ &= \hbar\sqrt{J(J+1) - M(M \mp 1)} \, \langle J, M \mp 1 | j_1, m_1; j_2, m_2 \rangle \ . \end{aligned} \tag{7.45}$$

Comparing the two sets of relations, we find that we have *identical recursion relations* with the relation

$$(J, M, j, m, k, q) \leftrightarrow (J, M, j_1, m_1, j_2, m_2) \ . \tag{7.46}$$

These recursion relations allow to determine both the Clebsch–Gordan coefficients (with fixed J, j_1, j_2) and our matrix elements (with fixed α, β, J, k, j) by homogeneous linear relations in terms of a single coefficient/matrix element. The two sets of numbers thus must agree up to an overall rescaling. ■

The Wigner–Eckart theorem tells us that the rules for addition of angular momentum also constrain the possible matrix elements of tensor operators in rotational systems. For example, for a vector operator like **X**, **P**, or even **J** itself, matrix elements can only be non-zero between states whose total angular momentum differs by at most one! This is an important selection rule in many applications, for example in atomic physics.

Chapter 8

Rayleigh–Schrödinger Perturbation Theory

In this chapter we turn to an important practical and conceptual question in quantum theory, which is how to approximate the energy levels of systems that are *close to* being described by a system that we can solve. We will say a little (but not a lot) more below about what we mean by "close", but intuitively, we can imagine (for example) modelling an atom with several electrons as being approximated by such an atom where the electrons do not interact among themselves except through their fermionic statistics. Alternatively, when we consider a Hydrogen-like, single-electron atom in the real world, we think of it as being well-described by the usual Hamiltonian with Coulomb potential, but there are actually small corrections to this due to the effects of (for instance) special relativity.

8.1 Formal perturbation theory

We imagine a scenario in which the Hamiltonian (for now we take this to be time-independent) of a quantum system can be written as a perturbation of a reference Hamiltonian $H^{(0)}$, so taking the form

$$H = H^{(0)} + \delta H \,. \tag{8.1}$$

We have in mind (though it need not be the case to develop the abstract formalism) that $H^{(0)}$ is a Hamiltonian whose spectral problem we have understood exactly, such as that of the harmonic oscillator or the Hydrogen atom (restricting to bound states). Now we want to consider δH as being small in an appropriate sense, and we make this explicit by writing

$$\delta H = \varepsilon H^{(1)} \,, \tag{8.2}$$

where ε represents a small parameter, either in the sense of being numerically small (with $H^{(1)}$ being somehow fixed in size and so not scaling like an inverse power of ε), or more accurately for our immediate purposes, in the sense of being *formally small*. So we will take our energy levels and stationary states to be formal series expansions in this parameter ε,

$$\begin{aligned}
\psi_\varepsilon &= \psi^{(0)} + \varepsilon\psi^{(1)} + \varepsilon^2\psi^{(2)} + \ldots + \varepsilon^n\psi^{(n)} + \ldots \,, \\
E_\varepsilon &= E^{(0)} + \varepsilon E^{(1)} + \varepsilon^2 E^{(2)} + \ldots + \varepsilon^n E^{(n)} + \ldots \,,
\end{aligned} \tag{8.3}$$

and ask that these satisfy the time-independent Schrödinger equation as formal series,

$$H\psi_\varepsilon = E_\varepsilon\psi_\varepsilon \,, \qquad \text{order by order in } \varepsilon \,. \tag{8.4}$$

The issue of normalisation for this formal solution can be a little bit subtle, since the normalisation of our formal solution will itself in principle be a formal series in ε. In particular, given one formal series solution of (8.4), we can produce another by multiplying both sides by (the same) formal series in ε with numerical coefficients.

It turns out that a convenient normalisation is, rather than demanding ψ_ε be unit normalised, to require

$$\left\langle \psi^{(0)} \middle| \psi_\varepsilon \right\rangle = 1 \,, \tag{8.5}$$

which is equivalent to

$$\left\langle \psi^{(0)} \middle| \psi^{(n)} \right\rangle = \delta_{n,0} \,. \tag{8.6}$$

It may be useful to ponder the enforcement of this normalisation if given a formal solution with an arbitrary normalisation.

Writing out (8.4) as an expansion in ε, we have,

$$\begin{aligned}\left(H^{(0)}+\varepsilon H^{(1)}\right)\left(\psi^{(0)}+\varepsilon\psi^{(1)}+\varepsilon^2\psi^{(2)}+\ldots\right) &= \\ \left(E^{(0)}+\varepsilon E^{(1)}+\varepsilon^2 E^{(2)}+\ldots\right)&\left(\psi^{(0)}+\varepsilon\psi^{(1)}+\varepsilon^2\psi^{(2)}+\ldots\right)\ ,\end{aligned} \tag{8.7}$$

and equating coefficients of order ε^n gives an infinite sequence of relations

$$H^{(0)}\psi^{(n)}+H^{(1)}\psi^{(n-1)}=\sum_{m=0}^{n}E^{(m)}\psi^{(n-m)}\ . \tag{8.8}$$

The first of these relations, where $n=0$ (by convention we let $\psi^{(-1)}=0$) just tells us that zeroth term in each expansion corresponds to an eigenstate and energy level of the unperturbed system,

$$H^{(0)}\psi^{(0)}=E^{(0)}\psi^{(0)}\ . \tag{8.9}$$

The higher order terms will then be our object of focus. However, first we should make a short comment about the interpretation of this formal series.

Remark 8.1.1 (Analytic considerations). The analytic status of these formal series solutions is an interesting subject. Natural questions include whether we can assign some numerical value to ε so that the resultant series expansions for the energies and the eigenstates converge, and if so what the radius of convergence might be. The following theorem due to Kato gives a sense of under what conditions we can guarantee convergence.

Theorem 8.1.2. If there exist real constants $a, b \geqslant 0$ such that for any ψ in the domain of $H^{(0)}$, we have

$$\|H^{(1)}\psi\|\leqslant a\|\psi\|+b\|H^{(0)}\psi\|\ , \tag{8.10}$$

then the formal series for ψ_ε and E_ε will have a non-zero radius of convergence.

Estimating this radius of convergence is a more subtle issue. Observe that for the case of a finite dimensional Hilbert space, this is always satisfied since all of the quantities appearing in (8.10) will be bounded. In the infinite dimensional case, the simplest scenario is again when the left hand side is simply bounded for all $\psi\in\mathcal{H}$; in such a case there is always a non-zero radius of convergence.
The cases that are most often of interest in a physical setting are of the more complicated variety, where $\mathcal{H}$ is infinite dimensional and the operators in question lead to unbounded left hand side of (8.10). In these cases establishing the bound in question might be a difficult problem. However, even when the series have zero radius of convergence, these expansions can often be interpreted as asymptotic series, and can be used to do computations that give good agreement with experiment (sometimes extraordinarily good agreement) by simply truncating the formal series after a fixed number of terms. In our applications this will be our *modus operandi*, and we will never have to grapple with the analytic subtleties of convergence.

8.2 First order perturbation theory (nondegenerate)

Now we return to our equation in (8.8) and look at the first correction term ($m=1$),

$$H^{(1)}\psi^{(0)}+H^{(0)}\psi^{(1)}=E^{(1)}\psi^{(0)}+E^{(0)}\psi^{(1)}\ . \tag{8.11}$$

Rewriting this gives

$$\left(H^{(0)}-E^{(0)}\right)\psi^{(1)}=-\left(H^{(1)}-E^{(1)}\right)\psi^{(0)}\ . \tag{8.12}$$

We want this equation to determine both $E^{(1)}$ and $\psi^{(1)}$, given the unperturbed (zeroth order) information. For now we assume that $E^{(0)}$ is a *nondegenerate energy level*, so $\psi^{(0)}$ is the only state vector (up to rescaling) with eigenvalue $E^{(0)}$. An important practical fact is that we can determine the energy correction $E^{(1)}$ without worrying about $\psi^{(1)}$. To do this, we take the inner product of both sides of equation (8.12) with the state vector $\psi^{(0)}$,

$$\left\langle \psi^{(0)} \middle| \left(H^{(0)} - E^{(0)}\right) \middle| \psi^{(1)} \right\rangle = - \left\langle \psi^{(0)} \middle| \left(H^{(1)} - E^{(1)}\right) \middle| \psi^{(0)} \right\rangle . \tag{8.13}$$

The left hand side vanishes, since $H^{(0)}$ can act to the left and just give $E^{(0)}$, and we get the equation for our energy correction

$$E^{(1)} = \left\langle \psi^{(0)} \middle| H^{(1)} \middle| \psi^{(0)} \right\rangle . \tag{8.14}$$

This simple expression is extremely important and extraordinarily powerful; in a slogan it tells us that the leading correction to the energy of a stationary state is given by the expectation value of the perturbation in that state.

We then turn to the problem of determining $\psi^{(1)}$. We will find an expression for the expansion of this correction vector in terms of a basis of stationary state vectors for the unperturbed Hamiltonian, which we denote by $\{\psi_n^{(0)}\}$, where

$$H^{(0)} \left| \psi_n^{(0)} \right\rangle = E_n^{(0)} \left| \psi_n^{(0)} \right\rangle . \tag{8.15}$$

We label these states so that the state whose perturbation we are studying is the m'th state ($n = m$), with energy $E_m^{(0)}$. We write such an expansion of the correction vector using a resolution of the identity,

$$\left| \psi_m^{(1)} \right\rangle = \sum_n \left| \psi_n^{(0)} \right\rangle \left\langle \psi_n^{(0)} \middle| \psi_m^{(1)} \right\rangle = \sum_{n \neq m} \left(\left\langle \psi_n^{(0)} \middle| \psi_m^{(1)} \right\rangle \right) \left| \psi_n^{(0)} \right\rangle . \tag{8.16}$$

where the $n = m$ term drops out of the summation due to our normalisation condition (8.5). Now we can compute the coefficients by taking the inner product of both sides of (8.12) with $\psi_n^{(0)}$, which yields

$$\begin{aligned}
\left\langle \psi_n^{(0)} \middle| \left(H^{(0)} - E_m^{(0)}\right) \middle| \psi_m^{(1)} \right\rangle &= - \left\langle \psi_n^{(0)} \middle| \left(H^{(1)} - E_m^{(1)}\right) \middle| \psi_m^{(0)} \right\rangle , \\
\implies \quad \left(E_n^{(0)} - E_m^{(0)}\right) \left\langle \psi_n^{(0)} \middle| \psi_m^{(1)} \right\rangle &= - \left\langle \psi_n^{(0)} \middle| H^{(1)} \middle| \psi_m^{(0)} \right\rangle , \\
\implies \quad \left\langle \psi_n^{(0)} \middle| \psi_m^{(1)} \right\rangle &= \frac{\langle \psi_n^{(0)} | H^{(1)} | \psi_m^{(0)} \rangle}{E_m^{(0)} - E_n^{(0)}} .
\end{aligned} \tag{8.17}$$

This allows us to write a general expression for the first correction to the mth energy eigenstate,

$$\psi_m^{(1)} = \sum_{n \neq m} \left(\frac{\langle \psi_n^{(0)} |\, H^{(1)} \,| \psi_m^{(0)} \rangle}{E_m^{(0)} - E_n^{(0)}} \right) \psi_n^{(0)} . \tag{8.18}$$

Remark 8.2.1. There is a nice way to understand this expression for the first correction to energy eigenstates that is somewhat slicker and provides a nice heuristic that we will reuse later. In principle, we would like to take (8.12) and solve for $\psi_m^{(1)}$ by *inverting* the operator appearing on the left hand side,

$$\left| \psi_m^{(1)} \right\rangle = \frac{-1}{H^{(0)} - E_m^{(0)}} \left(H^{(1)} - E_m^{(1)}\right) \left| \psi_m^{(0)} \right\rangle . \tag{8.19}$$

However, the operator $H^{(0)} - E_m^{(0)}$ is definitely *not* invertible, since it has a kernel (spanned by $\psi_m^{(0)}$ in this non-degenerate case) so is not injective. A related fact is that the range of this operator doesn't include all of $\mathcal{H}$; $\psi_m^{(0)}$ is absent, and indeed it is the only one of our basis vectors that is absent.
Consequently, we can only hope to define this inverse operator on the subspace of $\mathcal{H}$ that is orthogonal to $\psi_m^{(0)}$, and even then its action can only be defined only up to the possible addition of multiples of $\psi_m^{(0)}$. The first of these requirements gives us a solvability condition: the right hand side of (8.12) must lie in the orthogonal complement to $\psi^{(0)}$. This is precisely the condition that we used to determine $E_m^{(1)}$. The second issue means

that we must adopt a convention for the component of $\psi_m^{(1)}$ in the direction of $\psi_m^{(0)}$. This is exactly what our normalisation condition (8.5) accomplishes.

With both of those issues square away, we can get an expression for $\psi_m^{(1)}$ upon inserting a resolution of the identity for the orthogonal complement $\mathcal{H}^\perp$ to $\psi_m^{(0)}$. (This must act as the identity given that we have ensured that the state vector on the right of (8.12) is lies in $\mathcal{H}^\perp$.)

$$\begin{aligned} \left|\psi_m^{(1)}\right\rangle &= \frac{-1}{H^{(0)}-E_m^{(0)}}\sum_{n\neq m}\left|\psi_n^{(0)}\right\rangle\left\langle\psi_n^{(0)}\right|\left(H^{(1)}-E_m^{(1)}\right)\left|\psi_m^{(0)}\right\rangle\ , \\ &= \sum_{n\neq m}\left(\frac{\left\langle\psi_n^{(0)}\right|H^{(1)}\left|\psi_m^{(0)}\right\rangle}{E_m^{(0)}-E_n^{(0)}}\right)\left|\psi_n^{(0)}\right\rangle\ . \end{aligned} \tag{8.20}$$

In passing to the second line, we have defined the action of the inverted operator using the usual rule that on eigenvectors of an observable A with eigenvalue a, we let a function of the observable $f(A)$ act by $f(a)$. We observe that this (admittedly formal) procedure neatly reproduces the more piecemeal analysis carried out above.

Example 8.2.2 (Energy of the Helium atom ground state)**.** A classic application of first order perturbation theory is to the study of a di-electronic atom (an atom with two electrons) like Helium. For now we will ignore the effects of spin and fermionic statistics. We are therefore studying wavefunctions $\psi(\mathbf{x}_1,\mathbf{x}_2)$ of two position vectors, with a Hamiltonian of the form

$$H = H_1 + H_2 + H_{\text{int}}\ , \tag{8.21}$$

where

$$H_i = -\frac{\hbar^2}{2m}\nabla_i^2 - \frac{Zq_e^2}{|\mathbf{x}_i|}\ ,\ i=1,2, \qquad H_{\text{int}} = \frac{q_e^2}{|\mathbf{x}_1-\mathbf{x}_2|}\ . \tag{8.22}$$

Here the nuclear charge Z would be two for Helium, H_1 and H_2 are the standard single-electron Hamiltonians including a Coulomb potential from the nucleus, and H_{int} is the *interaction Hamiltonian* that encodes the repulsion between the two electrons.

If we could ignore the interaction term and had as our Hamiltonian just $H_1 + H_2$, then this system would be the tensor product of two copies of the Hydrogen atom system, and we would have stationary states that were separable,

$$\psi_{n_1,n_2}(\mathbf{x}_1,\mathbf{x}_2) = \psi_{n_1}(\mathbf{x}_1)\psi_{n_2}(\mathbf{x}_2)\ . \tag{8.23}$$

In particular, the ground state is non-degenerate and given by[51]

$$\begin{aligned} \psi_{1,1}(\mathbf{x}_1,\mathbf{x}_2) &= \left(\frac{Z^3}{\pi a_0^3}\right)^{\frac{1}{2}}\exp\left(-\frac{Zr_1}{a_0}\right)\times\left(\frac{Z^3}{\pi a_0^3}\right)^{\frac{1}{2}}\exp\left(-\frac{Zr_2}{a_0}\right)\ , \\ &= \left(\frac{Z^3}{\pi a_0^3}\right)\exp\left(-\frac{Z(r_1+r_2)}{a_0}\right)\ . \end{aligned} \tag{8.24}$$

Though it is certainly not clear that the interaction Hamiltonian can be thought of as being *small*, we can nevertheless proceed with a formal perturbative analysis setting $\delta H = H_{\text{int}}$.[52] The first order correction to the ground state energy using our machinery from above is then given by the expectation value,

$$E_1^{(1)} = \left\langle\psi^{(0)}\middle|H_{\text{int}}\middle|\psi^{(0)}\right\rangle = \left(\frac{Z^3}{\pi a_0^3}\right)^2\int_{\mathbb{R}^6}\frac{q_e^2\exp\left(\frac{-2Z(r_1+r_2)}{a_0}\right)}{|\mathbf{x}_1-\mathbf{x}_2|}\,\mathrm{d}^3\mathbf{x}_1\,\mathrm{d}^3\mathbf{x}_2\ . \tag{8.25}$$

Evaluating integrals like these is (unfortunately) largely unavoidable when it comes time to turn the abstract algebra of perturbation theory into actual numbers for systems like the Helium atom. We can perform this integral by first noting that spherical symmetry allows us to fix $\mathbf{x}_1$ to point in the $\hat{\mathbf{z}}$ direction (we then pick up a factor of 4π from the angular integration for $\mathbf{x}_1$). Introducing spherical polar coordinates for $\mathbf{x}_2$, we then have

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = r_1 r_2 \cos\theta \,, \qquad |\mathbf{x}_1 - \mathbf{x}_2| = (r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta)^{1/2} \,. \tag{8.26}$$

The φ integral for $\mathbf{x}_2$ is trivial and gives an extra factor of 2π. Thus the original six-dimensional integral reduces to

$$\left(\frac{8Z^6 q_e^2}{a_0^6}\right) \int_0^\infty \int_0^\infty \int_0^\pi \exp\left(\frac{-2Z(r_1+r_2)}{a_0}\right) \frac{r_1^2 r_2^2 \sin\theta}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2\cos\theta}}\, \mathrm{d}\theta\, \mathrm{d}r_1\, \mathrm{d}r_2 \,. \tag{8.27}$$

The angular integral can be done immediately,

$$\begin{aligned} \int_0^{2\pi} \frac{\sin\theta\, \mathrm{d}\theta}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2\cos\theta}} &= \frac{1}{r_1 r_2} \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta}\,\Bigg|_{\theta=0}^{\theta=\pi} \,. && (8.28) \\ &= \frac{1}{r_1 r_2} \left((r_1 + r_2) - |r_1 - r_2|\right) \,, && (8.29) \\ &= \begin{cases} \frac{2}{r_1} & r_1 > r_2 \,, \\ \frac{2}{r_2} & r_2 > r_1 \,. \end{cases} && (8.30) \end{aligned}$$

By symmetry the integration regions where $r_1 > r_2$ and that with $r_2 > r_1$ contribute equally, so we can write just take twice the part where $r_1 > r_2$ and the resulting radial integral becomes

$$E_1^{(1)} = \left(\frac{32Z^6 q_e^2}{a_0^6}\right) \int_0^\infty \int_{r_2}^\infty \exp\left(-\frac{2Z(r_1+r_2)}{a_0}\right) r_1 r_2^2\, \mathrm{d}r_1\, \mathrm{d}r_2 \,. \tag{8.31}$$

This leaves radial integrations, and these can be performed with the help of the following useful integral identity

Lemma 8.2.3. For any non-negative integer n, we have

$$\int_R^\infty e^{-kr} r^n\, \mathrm{d}r = \frac{n! e^{-kR}}{k^{n+1}} \sum_{j=0}^n \frac{(kR)^j}{j!} \,. \tag{8.32}$$

This is proved by induction on n by differentiating with respect to k.

Applying this lemme twice in (8.31) gives the final result,

$$E_1^{(1)} = \frac{5}{8} \frac{Z q_e^2}{a_0} \,. \tag{8.33}$$

Selecting $Z = 2$ as is appropriate for Helium and performing a crude truncation of the perturbation series to first order, this gives an estimation for the ground state energy as

$$E_1 \approx E_1^{(0)} + E_1^{(1)} = -\frac{2q_e^2}{a_0}\left(2 - \frac{5}{8}\right) = -\frac{2.75 q_e^2}{a_0} \,. \tag{8.34}$$

The experimental result for the Helium atom's ground state energy is

$$E_1^{(\mathrm{exp})} \approx -\frac{2.92 q_e^2}{a_0} \,. \tag{8.35}$$

Though we don't have systematic control of higher order corrections, it is gratifying to see that the estimate with a coefficient of 2.75 is much closer to the experimental value than the zeroth order estimate with a coefficient of 4.

8.3 First order perturbation theory (degenerate)

We need to be wary of the case where the energy level whose corrections we are considering is degenerate in the unperturbed theory. For example, we may be considering a Hydrogen-like ion where the energy levels all have some degeneracy. In this case, the analysis of the previous subsection needs to be improved a bit.

This is most readily observed by considering the calculation that yielded the first correction to the energy. Suppose that both $\psi^{(0)}$ and $\varphi^{(0)}$ both have unperturbed energy $E^{(0)}$. Taking the inner product of both sides of (8.12) with both $\psi^{(0)}$ and $\varphi^{(0)}$, we find

$$E^{(1)} = \left\langle \psi^{(0)} \middle| H^{(1)} \middle| \psi^{(0)} \right\rangle \,, \qquad 0 = \left\langle \varphi^{(0)} \middle| H^{(1)} \middle| \psi^{(0)} \right\rangle \,. \tag{8.36}$$

The first of these is the same relation we had before and could be thought to determine the first energy correction, but the second may or may not be satisfied depending on whether we have been lucky in our choice of basis vectors $\psi^{(0)}$ and $\varphi^{(0)}$. Indeed, the problem is precisely that there is an ambiguity in the basis that we choose for the degenerate $E^{(0)}$ eigenspace, whereas after perturbation there is a preferred choice of basis. This resolution proceeds as follows.

Theorem 8.3.1. Let $\varphi_1, \ldots, \varphi_d$ be an orthonormal basis for the $E^{(0)}$ eigenspace of $H^{(0)}$. Then (8.12) can be solved if and only if $E^{(1)}$ and $\psi^{(0)}$ are chosen so that $E^{(1)}$ is a solution to

$$\det\left(\left\langle \varphi_r \middle| H^{(1)} \middle| \varphi_s \right\rangle - E^{(1)}\delta_{rs}\right) = 0 \,, \tag{8.37}$$

i.e., $E^{(1)}$ should be an eigenvalue of the matrix whose with entries $\left\langle \varphi_r \middle| H^{(1)} \middle| \varphi_s \right\rangle$, and $\psi^{(0)} = \sum_{r=1}^{d} c_r \varphi_r$ is the corresponding eigenvector

$$\sum_s \left\langle \varphi_r \middle| H^{(1)} \middle| \varphi_s \right\rangle c_s = E^{(1)} c_r \,. \tag{8.38}$$

In a slogan, this says that before setting up the first order perturbation problem, one should choose a basis for the degenerate $E^{(0)}$ eigenspace that diagonalises the action of the restriction of $H^{(1)}$ to that subspace. In this basis, the rules are the same as in the non-degenerate setting.

Proof. This follows directly from the requirement that we avoid the potential contradictions arising as in (8.36), but we will argue more abstractly along the lines of Remark 8.2.1.

We need to be able to solve (8.12), but again $H^{(0)} - E^{(0)}$ has a (now d-dimensional) kernel and is not invertible, and we need to make our choices to ensure that the right hand side lies in the range of $H^{(0)} - E^{(0)}$. There is a helpful proposition whose proof will not be important, though we give most of it (omitting a technical detail).

Proposition 8.3.2. The range $\mathrm{Ran}(A)$ of a self-adjoint operator A on a Hilbert space $\mathcal{H}$ (for which zero is not an accumulation point of the spectrum) coincides with the orthogonal complement of its kernel.

Proof. We first establish that $\mathrm{Ran}(A) \subseteq \ker(A)^{\perp}$. Let $\psi \in \mathrm{Ran}(A)$, so we can write $\psi = A\psi'$. Then we have for any $\phi \in \ker(A)$,

$$\langle \phi | \psi \rangle = \langle \phi | A\psi' \rangle = \langle A\phi | \psi' \rangle = 0 \,, \tag{8.39}$$

which gives the inclusion. We next establish that $\mathrm{Ran}(A)^{\perp} \subseteq \ker(A)$. First, note that for $\phi \in \mathrm{Ran}(A)^{\perp}$, we have that for all $\psi \in \mathcal{H}$, we have

$$0 = \langle \phi | A\psi \rangle = \langle A\phi | \psi \rangle \,, \tag{8.40}$$

---

[51] In more detail, not ignoring spin, the ground state for a single electron is two dimensional and can be represented as $\psi_1(\mathbf{x})(\alpha\left|\tfrac{1}{2}\right\rangle + \beta\left|-\tfrac{1}{2}\right\rangle)$ with α, β some arbitrary complex numbers. With two particles in the ground state, we cannot make the wave function antisymmetric in the positions $\mathbf{x}_1$ and $\mathbf{x}_2$ because the ground state is unique, but we can make the total state antisymmetric using the spins, giving the two-particle ground state

$$|\psi\rangle = \psi_1(\mathbf{x}_1)\psi_1(\mathbf{x}_2)\left(\left|\tfrac{1}{2};-\tfrac{1}{2}\right\rangle - \left|-\tfrac{1}{2};\tfrac{1}{2}\right\rangle\right) \,,$$

where $\left|\tfrac{1}{2};-\tfrac{1}{2}\right\rangle$ is the state in which the first particle has intrinsic spin $m_s = \tfrac{1}{2}$ and the second $m_s = -\tfrac{1}{2}$, *etc.*. At the level of spins, or course, this is a simple case of addition of angular momentum where $\mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2} = \mathcal{H}_1 \oplus \mathcal{H}_0$ and we take the (unique) spin-0 state.

[52] One way one could contrive a more convincing setting for this perturbative analysis would be to take the large Z limit. Then the Coulomb potential due to the nucleus becomes arbitrarily strong compared to the inter-electronic interaction, so that the approximation of ignoring the interaction seems plausible.

so $A\phi$ must be the zero vector, and thus $\phi \in \ker(A)$, and taking orthogonal complements, $\ker(A)^\perp \subseteq (\mathrm{Ran}(A)^\perp)^\perp$. All that remains is make the identification[53]

$$(\mathrm{Ran}(A)^\perp)^\perp \overset{?}{=} \mathrm{Ran}(A)\ . \tag{8.41}$$

It turns out this is not true automatically; it requires the operator A to have a *closed range*. This, in turn, follows (via a straightforward but uninteresting-for-our-purposes proof) from the condition that zero not be an accumulation point in the spectrum of A. With both inclusions in place, we get then recover the stated identification. ■

Thus what we need to ensure is that the the right hand side of (8.12) is orthogonal to the operator's kernel. Taking the inner product with any one of the ϕ_r we find,

$$0 = \left\langle \phi_r \middle| \left(H^{(1)} - E^{(1)}\right) \middle| \psi^{(0)} \right\rangle = \sum_s \left\langle \phi_r \middle| H^{(1)} \middle| \phi_s \right\rangle c_s - E^{(1)} \delta_{rs} c_s\ , \tag{8.42}$$

where $\psi^{(0)}$ is expanded in terms of the ϕ_i as in the statement of the theorem. This is precisely the condition that $E^{(1)}$ be an eigenvalue of the matrix with matrix elements $\langle \phi_r | H^{(1)} | \phi_s \rangle$ with $\psi^{(0)}$ (thought of as an element of the d-dimensional $E^{(0)}$ eigenspace) the corresponding eigenvector. ■

In practice, what one should usually do is to choose the basis $\phi_1, \ldots, \phi_d$ to already diagonalise the restriction to the $E^{(0)}$ eigenspace of the perturbation $H^{(1)}$ (*i.e.*, to be the eigenvectors appearing in the above theorem). Using these as the unperturbed state vectors when seting up the formal perturbation problem (say, setting $\psi^{(0)} = \phi_1$) then gives for the first correction to such an eigenstate,

$$\psi^{(1)} \quad = \quad \sum_{E_n^{(0)} \neq E^{(0)}} \left(\frac{\langle \psi_n^{(0)} |\, H^{(1)} \,| \psi^{(0)} \rangle}{E^{(0)} - E_n^{(0)}} \right) \left| \psi_n^{(0)} \right\rangle + \sum_{j=2}^{d} \lambda_j \left| \phi_j \right\rangle\ , \tag{8.43}$$

where the parameters λ_i parameterise the ambiguity in defining the preimage of the right hand side of (8.12) due to the enlarged kernel of $H^{(0)} - E^{(0)}$. At this stage these parameters represent genuine ambiguities. You will see on the third problem sheet that these parameters are fixed upon continuing to higher orders in perturbation theory.

Example 8.3.3 (Strong field Zeeman effect). A simple example of degenerate perturbation theory arises in what is known as the Zeeman effect (or Zeeman splitting) for Hydrogen-like atoms. If we apply a constant magnetic field $\mathbf{B}$ to such an atom, there turns out to be a correction to the Hamiltonian of the following form,

$$\delta H = \frac{q_e}{2m_e c} \mathbf{B} \cdot (\mathbf{L} + g_e \mathbf{S})\ . \tag{8.44}$$

Here, $g_e \approx 2$ is a numerical constant known as the *gyromagnetic ratio* of the electron. Without loss of generality, we will take $\mathbf{B}$ to be oriented in the x_3-direction.

When considering the effect of this perturbation, we must account for the degeneracy of the Hydrogen energy levels, and as we understood in the previous chapter, there are at least two natural bases for the degenerate energy eigenspaces of the Hydrogen atom, depending on whether we want to diagonalise the total angular momentum ($|n, \ell, j, m_j\rangle$ states) or the x_3 component of both orbital and spin angular momentum ($|n, \ell, m_\ell; m_s\rangle$ states). For this perturbation, then, we should use the latter states since they are actually already simultaneous eigenvectors of L_3 and S_3.

If we restrict to states with principle quantum number n (so unperturbed energy E_n), then the first order corrections are given by

$$E^{(1)}_{n,\ell,m_\ell,m_s} = \langle n, \ell, m_\ell; m_s | \delta H | n, \ell, m_\ell; m_s \rangle = \frac{q_e B}{2 m_e c} (m_\ell + g_e m_s)\ . \tag{8.45}$$

---

[53] It should be a familiar fact that the iterated orthogonal complement is the identity for finite-dimensional inner product spaces; the novelty here is that in infinite dimensions it isn't necessarily true.

The above example is a bit artificial as an example of perturbation theory, because the $|n,\ell,m_\ell;m_s\rangle$ states actually are *exact eigenstates* of the perturbed Hamiltonian, as they diagonalise the J_3 and S_3 operators appearing in δH on the nose. Correspondingly, you should be able to convince yourself that the correction (8.43) to the energy eigenstates vanishes exactly. Therefore this analysis is valid for *large* values of B. In fact, it is *only* value for sufficiently large values of B, because for small B there are competing corrections that need to be accounted for.

Example 8.3.4 (Fine structure of Hydrogen)**.** More elaborate applications of degenerate perturbation theory arise in a relativistic account of the Hydrogen atom. The relativistic treatment of the electron requires replacing the usual Schrödinger equation with something known as the *Dirac equation*, but in the context of the Hydrogen atom, the corrections due to relativistic effects are quite small, and can be treated perturbatively. A complete treatment of the first order relativistic corrections to the Hydrogen atom can be found in many textbooks, but for our purposes here we will restrict to just one (of three) correction term(s)—the spin orbit coupling:

$$\delta H_{SO} = \frac{q_e^2}{2m_e^2c^2}\frac{\mathbf{L}\cdot\mathbf{S}}{r^3}\,. \tag{8.46}$$

Intuitively, this term arises from the interaction between the spin of the electron and and the magnetic field it experiences in its rest frame due to its motion through the electric field of the nucleus.

According to our previous discussion of deegenerate first order perturbation theory, our task is to find the eigenvalues of $\langle\varphi_r|\delta H_{SO}|\varphi_s\rangle$ where the $\{\varphi_r\}$ span an eigenspace $E_n^{(0}$ of $H^{(0)}$. The eigenspaces of $H^{(0)}$ are highly degenerate, being determined by just the principal quantum number n, and we previously encountered two standard bases for these states: $\{|n,\ell,m_\ell;m_s\rangle\}$ and $\{|n,\ell,j,m_j\rangle\}$. It turns out that the restriction of the spin-orbit perturbation, $\langle\varphi_r|H'|\varphi_s\rangle$, is already diagonal in the added-spin basis $\{|n,\ell,j,m_j\rangle\}$. This is because we can write

$$\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}(\mathbf{J}^2-\mathbf{L}^2-\mathbf{S}^2)\,, \tag{8.47}$$

and so δH_{SO}, $\mathbf{J}^2$, $\mathbf{L}^2$, $\mathbf{S}^2$, and J_3 all commute. Thus we have

$$\left\langle n,\ell,j,m_j\left|\frac{\mathbf{L}\cdot\mathbf{S}}{r^3}\right|n,\ell',j',m_j'\right\rangle = \delta_{\ell\ell'}\delta_{jj'}\delta_{m_jm_j'}\frac{\hbar^2}{2}\left(j(j+1)-\ell(\ell+1)-s(s+1)\right)\left\langle\frac{1}{r^3}\right\rangle_{n,\ell}\,, \tag{8.48}$$

where $s=1/2$ for the electron and the expectation value can be computed just using radial wavefunctions. We quote that, for the n^{th} radial eigenfunction,

$$\left\langle\frac{1}{r^3}\right\rangle_{n,\ell} = \int_0^\infty \mathrm{d}r\,\frac{|f_{n,\ell}(r)|^2}{r^3}r^2 = \frac{1}{a_0^3n^3\ell(\ell+\frac{1}{2})(\ell+1)}\,, \tag{8.49}$$

where as usual, $a_0=\hbar^2/(mq_e^2)$ is the Bohr radius. Thus, for the full first-order correction from the spin-orbit term we have

$$\begin{aligned} E^{(1)}_{n,\ell,j} &= \frac{q_e^4}{4a_0^2n^3m_ec^2}\frac{(j(j+1)-\ell(\ell+1)-\frac{3}{4})}{\ell(\ell+\frac{1}{2})(\ell+1)}\,,\\ &= \frac{n\left(E_n^{(0)}\right)^2}{m_ec^2}\left(\frac{j(j+1)-\ell(\ell+1)-\frac{3}{4}}{\ell(\ell+\frac{1}{2})(\ell+1)}\right)\,. \end{aligned} \tag{8.50}$$

Recalling that $j=\ell\pm\frac{1}{2}$ for the spin one-half electron, this simplifies to

$$E^{(1)}_{n,\ell,j} = \begin{cases} \frac{n\left(E_n^{(0)}\right)^2}{m_ec^2}\left(\frac{1}{(\ell+\frac{1}{2})(\ell+1)}\right)\,, & j=\ell+\frac{1}{2}\\ \frac{n\left(E_n^{(0)}\right)^2}{m_ec^2}\left(\frac{-1}{\ell(\ell+\frac{1}{2})}\right)\,, & j=\ell-\frac{1}{2} \end{cases}\,. \tag{8.51}$$

The above manipulation actually breaks down for $\ell=0$, where the radial integral (8.49) is divergent. In the case of an $\ell = 0$ state, the spin orbit coupling should be thought of as vanishing identically (roughly, because **L** acts as zero). But surprisingly, the above expression (8.51) actually applies correctly upon setting $\ell = 0$, noting that the divergence has been cancelled at this point and only the first case in the equation obtains. However, to compute the $\ell = 0$ correction and confirm this assertion requires including another correction term (the *Darwin term*) in the Hydrogen Hamiltonian.

8.4 Higher order perturbation theory

We will briefly comment on the extension of this approach to higher orders in the perturbative expansion. To begin, let us just proceed directly to second order, where the $O(\varepsilon^2)$ term in (8.7) amounts to the equation

$$H^{(0)}\psi^{(2)} + H^{(1)}\psi^{(1)} = E^{(0)}\psi^{(2)} + E^{(1)}\psi^{(1)} + E^{(2)}\psi^{(0)} . \tag{8.52}$$

We reorganise this suggestively as

$$\left(H^{(0)} - E^{(0)}\right)\psi^{(2)} = -\left(H^{(1)} - E^{(1)}\right)\psi^{(1)} + E^{(2)}\psi^{(0)} , \tag{8.53}$$

which again tells us that to be able to determine $\psi^{(2)}$ we have to find the preimage under $H^{(0)} - E^{(0)}$ of the vector on the right hand side. If for now we assume that the unperturbed energy level in question is non-degenerate, then as at first order, our first task will be to establish a solvability condition that the right hand side is orthogonal to the kernel of $H^{(0)} - E^{(0)}$, which is to say that it is orthogonal to $\psi^{(0)}$. This determines $E^{(2)}$:

$$E^{(2)} = \left\langle \psi^{(0)} \middle| \left(H^{(1)} - E^{(1)}\right) \middle| \psi^{1} \right\rangle = \left\langle \psi^{(0)} \middle| H^{(1)} \middle| \psi^{1} \right\rangle , \tag{8.54}$$

which, plugging in our expression for $\psi^{(1)}$, gives

$$E^{(2)} = \sum_{\psi_n^{(0)} \neq \psi^{(0)}} \frac{\langle \psi^{(0)} | H^{(1)} | \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | H^{(1)} | \psi^{(0)} \rangle}{E^{(0)} - E_n^{(0)}} = \sum_{\psi_n^{(0)} \neq \psi^{(0)}} \frac{\left| \langle \psi_n^{(0)} | H^{(1)} | \psi^{(0)} \rangle \right|^2}{E^{(0)} - E_n^{(0)}} . \tag{8.55}$$

Let us make a few comments.

- If $\psi^{(0)}$ is the unperturbed ground state, then the denominator in (8.55) is always negative and so the second order energy correction is necessarily negative.

- This formula is especially useful if $H^{(1)}\psi^{(0)}$ is itself an energy eigenstate, so that we only get one nonvanishing term in the sum due to the orthogonality of the different energy eigenstates.

- If energy levels are well separated, then we expect the terms corresponding to mixing with nearby energy eigenstates to dominate the correction due to suppression in the denominator by energy difference.

Continuing to the second-order correction to the energy eigenstate (subject to our normalisation condition), we will have an expansion,

$$\psi^{(2)} = \sum_{\psi_n^{(0)} \neq \psi^{(0)}} \left\langle \psi_n^{(0)} \middle| \psi^{(2)} \right\rangle \psi_n^{(0)} , \tag{8.56}$$

and we can solve for the coefficients using (8.52). Ultimately, we find[54]

$$\begin{aligned}\psi^{(2)} = \sum_{\psi^{(0)}_{n,m}\neq\psi^{(0)}} \frac{\left\langle \psi^{(0)}_n \middle| H^{(1)} \middle| \psi^{(0)}_m \right\rangle \left\langle \psi^{(0)}_m \middle| H^{(1)} \middle| \psi^{(0)} \right\rangle}{\left(E^{(0)}_n - E^{(0)}\right)\left(E^{(0)}_m - E^{(0)}\right)} \psi^{(0)}_n - \\ \sum_{\psi^{(0)}_n\neq\psi^{(0)}} \frac{\left\langle \psi^{(0)}_n \middle| H^{(1)} \middle| \psi^{(0)} \right\rangle \left\langle \psi^{(0)} \middle| H^{(1)} \middle| \psi^{(0)} \right\rangle}{\left(E^{(0)}_n - E^{(0)}\right)^2} \psi^{(0)}_n \ .\end{aligned} \tag{8.58}$$

Clearly, at this point (and even more at higher orders) things get quite cluttered in these general expressions.

Often it is more important to have expressions for higher order corrections to energy levels than it is to have the high order expressions for the actual state vectors. From our general perturbative expansion, this would naively require obtaining $\psi^{(k-1)}$ if we want to compute $E^{(k)}$ because expanding $H_\varepsilon\psi_\varepsilon = E_\varepsilon\psi_\varepsilon$ to order k gives from the coefficient of ε^k

$$H^{(0)}\psi^{(k)} + H^{(1)}\psi^{(k-1)} = E^{(0)}\psi^{(k)} + E^{(1)}\psi^{(k-1)} + \ldots + E^{(k)}\psi^{(0)} \ , \tag{8.59}$$

and taking the inner product with $\psi^{(0)}$ (imposing the solvability condition in the nondegenerate case), using $\left\langle \psi^{(0)} \middle| \psi^{(l)} \right\rangle = 0$ that follows from the normalization condition, we obtain

$$E^{(k)} = \left\langle \psi^{(0)} \middle| H^{(1)} \middle| \psi^{(k-1)} \right\rangle \ , \tag{8.60}$$

In fact, one can avoid calculating all the way up to $\psi^{(k-1)}$ using a clever trick that draws out relations between the different perturbative corrections.

Lemma 8.4.1. As a formal expansion in ε_1 and ε_2, we have

$$\left\langle \psi_{\varepsilon_1} \middle| H^{(1)} \middle| \psi_{\varepsilon_2} \right\rangle = \frac{E_{\varepsilon_1} - E_{\varepsilon_2}}{\varepsilon_1 - \varepsilon_2} \left\langle \psi_{\varepsilon_1} \middle| \psi_{\varepsilon_2} \right\rangle \ . \tag{8.61}$$

Proof. We have $H_{\varepsilon_1} - H_{\varepsilon_2} = (\varepsilon_1 - \varepsilon_2)H^{(1)}$ so we can compute,

$$\left\langle \psi_{\varepsilon_1} \middle| H_{\varepsilon_1} - H_{\varepsilon_2} \middle| \psi_{\varepsilon_2} \right\rangle = \left\langle \psi_{\varepsilon_1} \middle| H_{\varepsilon_1} \middle| \psi_{\varepsilon_2} \right\rangle - \left\langle \psi_{\varepsilon_1} \middle| H_{\varepsilon_2} \middle| \psi_{\varepsilon_2} \right\rangle \ , \tag{8.62}$$

$$(\varepsilon_1 - \varepsilon_2)\left\langle \psi_{\varepsilon_1} \middle| H^{(1)} \middle| \psi_{\varepsilon_2} \right\rangle = E_{\varepsilon_1}\left\langle \psi_{\varepsilon_1} \middle| \psi_{\varepsilon_2} \right\rangle - E_{\varepsilon_2}\left\langle \psi_{\varepsilon_1} \middle| \psi_{\varepsilon_2} \right\rangle \ . \tag{8.63}$$

which leads to the conclusion. ■

Note that $E_{\varepsilon_1} - E_{\varepsilon_2}$ is proportional to $\varepsilon_1 - \varepsilon_2$, so the right hand side can still be taken as a formal expansion in the two variables.

By expanding both sides of (8.61), and considering each monomial in the ε's, we generate various identities for the perturbative corrections. Indeed, our first naive higher order formula (8.60) represents the coefficient of ε_2^{k-1}. However, we can be a bit more efficient if we are clever.

Corollary 8.4.2. $E^{(2k+1)}$ can be expressed in terms of only the $\psi^{(r)}$ with $r \leqslant k$ and and their matrix elements with respect to $H^{(1)}$.

---

[54] An alternate way to organise this calculation is to write

$$\psi^{(2)} = -\frac{1}{\left(H^{(0)} - E^{(0)}\right)} \Pi_{\mathcal{H}^{\perp}_{E^{(0)}}} \left(H^{(1)} - E^{(1)}\right)\psi^{(1)} \ , \tag{8.57}$$

where the projection operator $\Pi_{\mathcal{H}^{\perp}_{E^{(0)}}}$ serves to eliminate components in the direction of $\psi^{(0)}$, which is the result of appropriately fixing the value of $E^{(2)}$.

Proof. The coefficient on the right hand side of (8.61) can be expanded as

$$\frac{E_{\varepsilon_1} - E_{\varepsilon_2}}{\varepsilon_1 - \varepsilon_2} = \sum_{n=1}^{\infty} \frac{{\varepsilon_1}^n - {\varepsilon_2}^n}{\varepsilon_1 - \varepsilon_2} E^{(n)} \tag{8.64}$$

$$= \sum_{n=1}^{\infty} E^{(n)} \left(\sum_{j=0}^{n-1} {\varepsilon_1}^j {\varepsilon_2}^{n-1-j} \right) , \tag{8.65}$$

so at order ${\varepsilon_1}^k {\varepsilon_2}^k$ we have $E^{(2k+1)}$. This is then equal to the coefficient of ${\varepsilon_1}^k {\varepsilon_2}^k$ in the expansion of

$$\frac{\left\langle \psi_{\varepsilon_1} \middle| H^{(1)} \middle| \psi_{\varepsilon_2} \right\rangle}{\left\langle \psi_{\varepsilon_1} \middle| \psi_{\varepsilon_2} \right\rangle} . \tag{8.66}$$

That term clearly depends only on $\psi^{(r)}$ for $r \leqslant k$. ■

The simplest example of this gives us an economical expression for the third order energy correction,

$$E^{(3)} = \left\langle \psi^{(1)} \middle| H^{(1)} \middle| \psi^{(1)} \right\rangle - E^{(1)} \left\langle \psi^{(1)} \middle| \psi^{(1)} \right\rangle , \tag{8.67}$$

which we can calculate using only the results from our previous first-order computations.

Chapter 9

Variational Methods

It is no surprise that many (or most) quantum systems of practical interest are not particularly close to being exactly solvable, rendering perturbation theory at least difficult to control and at worst completely unhelpful. There are an important class of techniques that go by the name of *variational methods* that can be applied quite generally in these circumstances, though they often require a bit of creativity to exploit well. In this chapter we present the essential idea behind these methods and study (again!) the Helium atom as a nice example.

9.1 Rayleigh quotients for observables

The main tool behind the variational methods of this chapter will be the so-called *Rayleigh quotient*.

Definition 9.1.1 (Rayleigh Quotient). For an observable A, the (real) function

$$\begin{aligned} f_A : \mathcal{H} &\longrightarrow \mathbb{R}\,, \\ \psi &\longmapsto \mathbb{E}_\psi(A) = \frac{\langle\psi|A|\psi\rangle}{\langle\psi|\psi\rangle}\,. \end{aligned} \tag{9.1}$$

is known as the *Rayleigh quotient* for A.

It may seem a bit overwrought to introduce a new name for what we already know as the expectation value of A in the state ψ. The point here is to highlight the fact that this is now being thought of as a (smooth) function on $\mathcal{H}$, which may not have been the way you previously thought of expectation values.

A key property of the Rayleigh quotient is that the stationary values/vectors for $f_A(\psi)$ are precisely eigenvalue/vectors for A. This can be demonstrated by first establishing the following technical lemma.

Lemma 9.1.2. Given a linear subspace $\mathcal{K} \subseteq \mathcal{H}$, then

$$\frac{\mathrm{d}}{\mathrm{d}t} f_A(\psi + t\phi)\bigg|_{t=0} = 0\,, \quad \forall \phi \in \mathcal{K} \quad \Longleftrightarrow \quad (A - f_A(\psi))\,|\psi\rangle \in \mathcal{K}^\perp\,. \tag{9.2}$$

Stated differently, this means that $\psi \in \mathcal{K}$ is a stationary vector of the Rayleigh quotient if and only if it is an eigenvector of the restricted Hamiltonian $\Pi_\mathcal{K} H \Pi_\mathcal{K} : \mathcal{K} \to \mathcal{K}$.

Proof. For $\phi \in \mathcal{K}$, also $i\phi \in \mathcal{K}$, so stationarity with respect to adding ϕ implies

$$\frac{\mathrm{d}}{\mathrm{d}t} f_A(\psi + t\phi) = 0 \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}t} f_A(\psi + it\phi) = 0\,. \tag{9.3}$$

By direct computation, the first of these relations gives

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t} \frac{\langle\psi + t\phi|A|\psi + t\phi\rangle}{\langle\psi + t\phi|\psi + t\phi\rangle}\bigg|_{t=0}\,, && (9.4) \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \frac{\langle\psi|A|\psi\rangle + t\,\langle\phi|A|\psi\rangle + t\,\langle\psi|A|\phi\rangle + O(t^2)}{\langle\psi|\psi\rangle + t\,\langle\phi|\psi\rangle + t\,\langle\psi|\phi\rangle + O(t^2)}\bigg|_{t=0}\,, && (9.5) \\ &= \frac{\langle\phi|A|\psi\rangle + \langle\psi|A|\phi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\psi|A|\psi\rangle\,(\langle\phi|\psi\rangle + \langle\psi|\phi\rangle)}{\langle\psi|\psi\rangle^2}\,, && (9.6) \\ &= 2\Re\left(\frac{\langle\phi|(A - f_A(\psi))|\psi\rangle}{\langle\psi|\psi\rangle}\right)\,. && (9.7) \end{aligned}$$

This gives the real part of the desired equation. If we repeat the argument with $\phi \to i\phi$ we get the imaginary part as well, and thus deduce a state being stationary under the addition of vectors $\phi \in \mathcal{K}$ is equivalent to it satisfying

$$\langle \phi | (A - f_A(\psi)) | \psi \rangle = 0 \,, \tag{9.8}$$

so equivalently, $(A - f_A(\psi))\,|\psi\rangle \in \mathcal{K}^\perp$. ■

If we set $\mathcal{K} = \mathcal{H}$, then $\mathcal{K}^\perp = \{0\}$ and we have the strong relation

$$(A - f_A(\psi))\,|\psi\rangle = 0 \,. \tag{9.9}$$

This means that the critical points of $f_A(\psi)$ are precisely the eigenvectors of A and the critical values are the eigenvalues of A.

Remark 9.1.3. Usually our observables are unbounded operators, so we can't really choose $\mathcal{K} = \mathcal{H}$, rather we can choose $\mathcal{K} = D(A)$, the domain of A. Fortunately, the conclusion still applies as long as $D(A)$ is a dense in $\mathcal{H}$, as the orthocomplement of a dense subset of a Hilbert space is empty.

9.2 The virial theorem

With the stationarity property we just established for the the Rayleigh quotient, we can obtain useful—and quite general—information about energy eigenstates by considering variations within one-parameter families of states. An important example is the following important theorem, which generalises an analogous result in classical mechanics.

Theorem 9.2.1 (Virial theorem)**.** Assume the Hamiltonian for a quantum mechanical system has the conventional form $H = T + V$ with

$$T = \frac{\mathbf{P}^2}{2m} = -\frac{\hbar^2}{2m}\nabla^2 \,, \qquad V = V(\mathbf{X}) \,. \tag{9.10}$$

Then for any stationary state ψ (so $H\psi = E\psi$), the following condition holds:

$$2\mathbb{E}_\psi(T) = \mathbb{E}_\psi(\mathbf{x} \cdot \nabla V) \,. \tag{9.11}$$

If V is homogeneous of degree N (*i.e.*, $V(\lambda\mathbf{x}) = \lambda^N V(\mathbf{x})$), then we have the following stronger result,

$$\mathbb{E}_\psi(T) = \frac{N}{N+2}E \,, \qquad \mathbb{E}_\psi(V) = \frac{2}{N+2}E \,. \tag{9.12}$$

Proof. The idea of the proof is to perform a variational analysis for the family of wave functions of the form $\psi_\lambda(\mathbf{x}) = \lambda^{d/2}\psi(\lambda\mathbf{x})$ for a given reference wave function ψ. The factor of $\lambda^{d/2}$ is included to ensure that all of these wave functions have the same normalisation; this is a matter of convenience, which allows us to ignore the denominator when computing Rayleigh quotients.

Now suppose that ψ is a (normalised) stationary state; then the Rayleigh quotient of ψ_λ must be stationary as a function of λ at $\lambda = 1$,[55]

$$\left.\frac{\mathrm{d}}{\mathrm{d}\lambda} f_H(\psi_\lambda)\right|_{\lambda=1} = 0 \,. \tag{9.13}$$

Proceeding by direct computation, we have by the chain rule $\nabla\psi_\lambda(\mathbf{x}) = \lambda^{\frac{d+2}{2}}\left(\nabla\psi\right)(\lambda\mathbf{x})$, which gives us for the

---

[55] Note that while Lemma (9.1.2) is formulated for linear subspaces of $\mathcal{H}$, the conclusion implies that in any smoothly parameterised family of states $\{\psi_{\vec{\lambda}}\}$ that includes a stationary state, that stationary states must give a points of $f_H(\vec{\lambda})$.

Rayleigh quotient,

$$
\begin{aligned}
f_H(\psi_\lambda) &= \int_{\mathbb{R}^d} \left(\frac{-\hbar^2\lambda^{d+2}}{2m} |(\nabla\psi)(\lambda\mathbf{x})|^2 + \lambda^d V(\mathbf{x})|\psi(\lambda\mathbf{x})|^2 \right) \mathrm{d}^d\mathbf{x}\,, \\
&= \int_{\mathbb{R}^d} \left(\frac{-\lambda^2\hbar^2}{2m} |\nabla'\psi(\mathbf{x}')| + V(\lambda^{-1}\mathbf{x}')|\psi(\mathbf{x}')|^2 \right) \mathrm{d}^d\mathbf{x}'\,, \\
&= \lambda^2\mathbb{E}_\psi(T) + \mathbb{E}_\psi(V(\lambda^{-1}\mathbf{x}))\,.
\end{aligned}
\tag{9.14}
$$

In the passing to the second line we have defined $\mathbf{x}' = \lambda\mathbf{x}$. With this, (9.13) gives the condition

$$
0 = 2\mathbb{E}_\psi(T) + \frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbb{E}_\psi(V(\lambda^{-1}\mathbf{x}))\bigg|_{\lambda=1} = 2\mathbb{E}_\psi(T) - \mathbb{E}_\psi(\mathbf{x}\cdot\nabla V(\mathbf{x}))\,,
\tag{9.15}
$$

which reproduces (9.11).

For V homogeneous of degree N, $\mathbf{x}\cdot\nabla V = NV$ and we get the simpler result

$$
2\mathbb{E}_\psi(T) = N\mathbb{E}_\psi\left(V\right)\,.
\tag{9.16}
$$

We also have for an energy eigenstate

$$
E = \mathbb{E}_\psi(T) + \mathbb{E}_\psi(V)\,,
\tag{9.17}
$$

and putting these together gives

$$
\mathbb{E}_\psi(T) = \frac{N}{N+2}E\,, \qquad \mathbb{E}_\psi(V) = \frac{2}{N+2}E\,,
\tag{9.18}
$$

as required. ■

Remark 9.2.2. We offer some additional comments here.

- We see that for the Coulomb potential, for which $N = -1$, we must have $E < 0$ because $V < 0$ and $\mathbb{E}_\psi(V)$ is twice the size of $\mathbb{E}_\psi(T)$. (Alternatively, T is a positive operator, so the fact that $\mathbb{E}_\psi(T) = -E$ gives the negativity of E.)
- For the harmonic oscillator, kinetic and potential energies are equal, $\mathbb{E}_\psi(V) = \mathbb{E}_\psi(T) = \frac{1}{2}E$.
- This *quantum virial theorem* has a classical version, which is the original virial theorem. The classical theorem has takes much the same form, but with expectation values of quantum observables replaced by *time averages* of the classical counterparts over classical trajectories.

9.3 Approximating the ground state

The flagship application of variational methods in quantum theory is to the study of the ground state and ground state energy of complex quantum systems. To this end, one observes the following.

Proposition 9.3.1. If f_H is bounded below and achieves its minimum $E_0 := \inf_{\mathbb{P}(\mathcal{H})} f_H$, then E_0 is the ground state energy (minimum eigenvalue) and any state ψ for which $f_H(\psi) = E_0$ is a ground state.

Proof. f_H is stationary at its minimum, so this will necessarily correspond to an eigenstate. Its eigenvalue will be the minimal one because all other eigenvalues are also realised as values of f_H. ■

Conversely, when a system does have a normalisable ground state ψ_0, the function f_H achieves its lower bound at ψ_0.

This relatively simple observation is very powerful. What it suggests is that we can try to find an approximate ground state by finding the minimum $E_{0,\text{approx}}$ of f_H restricted to some cleverly chosen subset of $\mathcal{H}$. We are guaranteed that $E_{0,\text{approx}} \geqslant E_0$ by the above, so this procedure produces rigorous upper bounds for E_0. If we are lucky and creative, such an upper bound may also be good *approximations* (though to estimate errors would be beyond the scope of the discussion here).

Example 9.3.2 (Helium again). We return to the two-electron Helium atom, with Hamiltonian

$$H = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_2^2}{2m} - 2q_e^2\left(\frac{1}{|\mathbf{x}_1|} + \frac{1}{|\mathbf{x}_2|}\right) + \frac{q_e^2}{|\mathbf{x}_1 - \mathbf{x}_2|} = T + V + \delta H\,, \tag{9.19}$$

where

$$T = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_2^2}{2m}\,, \quad V = -2q_e^2\left(\frac{1}{|\mathbf{x}_1|} + \frac{1}{|\mathbf{x}_2|}\right)\,, \quad \delta H = \frac{q_e^2}{|\mathbf{x}_1 - \mathbf{x}_2|}\,. \tag{9.20}$$

As we saw in the previous chapter, we can try to treat δH as a small correction and use first-order perturbation theory to get *decent* results for the ground state energy. However, the approximation in that case was theoretically quite uncontrolled. This is a perfect case in which to attempt a variational estimate, and indeed we can do quite well.

We will use a one-parameter family of wavefunctions like we did with the virial theorem. To motivate our *variational Ansatz*, we use a physical argument. The idea is that if we want to model the dynamics of two electrons in the Helium atom as being non-interacting, we should adjust the effective value of the nuclear charge to account for screening, *i.e.*, each electron should on average see less than the full charge of the nucleus since the other electron is producing an electric field with the opposite sign. In other circumstances, this kind of an approximation is sometimes called a *mean field approximation*, and it can be quite effective.

To put this idea into practice, we adopt trial wave functions of the form

$$\psi_Z(\mathbf{x}_1, \mathbf{x}_2) = \left(\frac{Z^3}{\pi a_0^3}\right)\exp\left(-\frac{Z(r_1 + r_2)}{a_0}\right)\,, \tag{9.21}$$

where Z is now a variational parameter with respect to which we will minimise. This is the exact ground state for the “effective Hamiltonian”,

$$H_Z = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_2^2}{2m} - Zq_e^2\left(\frac{1}{r_1} + \frac{1}{r_2}\right) = T + \frac{Z}{2}V\,, \tag{9.22}$$

satisfying $H_Z\psi_Z = E_Z\psi_Z$ with $E_Z = -Z^2q_e^2/a_0$.

We now want to minimise the Rayleigh quotient associated to the true Hamiltonian H evaluated on these states as a function of Z, so we need to compute

$$f_H(\psi_Z) = \mathbb{E}_{\psi_Z}(T) + \frac{2}{Z}\mathbb{E}_{\psi_Z}\left(\frac{Z}{2}V\right) + \mathbb{E}_{\psi_Z}(\delta H)\ . \tag{9.23}$$

We have multiplied and divided by $2/Z$ in the second term in order to make it claer that the first two terms can be evaluated using the virial theorem with respect to the effective Hamiltonian H_Z. Indeed, we have

$$\mathbb{E}_{\psi_Z}(T) = -E_Z = \frac{Z^2q_e^2}{a_0}\,, \qquad \mathbb{E}_{\psi_Z}\left(\frac{Z}{2}V\right) = 2E_Z = -\frac{2Z^2q_e^2}{a_0}\,. \tag{9.24}$$

The last term in (9.23) is the matrix element that what we computed previously when we analysed this problem using first-order perturbation theory, and we import the result here,

$$\langle\psi_Z|\delta H|\psi_Z\rangle = q_e^2\left\langle\psi_Z\middle|\frac{1}{|\mathbf{x}_1 - \mathbf{x}_2|}\middle|\psi_Z\right\rangle = \frac{5}{8}\frac{Zq_e^2}{a_0}\,. \tag{9.25}$$

Putting it all together, we have for our Rayleigh quotient

$$f_H\left(\psi_Z\right) = \frac{q_e^2}{a_0}\left(Z^2 - \frac{27}{8}Z\right) = \frac{q_e^2}{a_0}\left[\left(Z - \frac{27}{16}\right)^2 - \left(\frac{27}{16}\right)^2\right] . \tag{9.26}$$

As a function of Z, this is minimised at $Z = \frac{27}{16}$, which we can interpret as telling us the extent to which each electron shields the charge of the nucleus from the perspective of the other electron. We interpret the minimal value as an upper bound on the Helium ground state energy,

$$E_0 \leqslant f_H(\psi_{\frac{27}{16}}) = -\frac{q_e^2}{a_0}\left(\frac{27}{16}\right)^2 \approx 2.85\frac{q_e^2}{a_0} . \tag{9.27}$$

This should be compared to our estimate of $-2.75q_e^2/a_0$ from first-order perturbation theory and the experimental value of approximately $-2.92q_e^2/a_0$; the variational upper bound is well below the first order estimate, and indeed is within three percent of the experimental value.
The variational estimate for the Helium ground state energy can be substantially improved by considering a more general Ansatz for the trial wave function. Indeed, an accuracy to within three parts in 10^8 has been obtained for this calculation using a family of *393 basis functions*.

The above result represents a marked improvement over what found using elementary perturbation theory. The fact that we do better is not a coincidence at all. Indeed, we have the following.

Proposition 9.3.3. Let $H = H^{(0)} + \delta H$ as before and let the family of states $\{\psi_\lambda\}$ over which we will minimise be chosen so that $\psi_0^{(0)}$ (the ground state of $H^{(0)}$) is contained in the family. Let $E_0^{(0)} + E_0^{(1)}$ be the first order perturbation theoretic estimate of the ground state energy found before. Then for $E_{\text{var}} = \inf_{\{\psi_\lambda\}} f_H$ we have

$$E_0^{(0)} + E_0^{(1)} \geqslant E_{\text{var}} \geqslant E_{\text{ground}} . \tag{9.28}$$

Proof. Recall that $E_0^{(1)} = \left\langle \psi_0^{(0)} \middle| \delta H \middle| \psi_0^{(0)} \right\rangle$ so that

$$E_0^{(0)} + E_0^{(1)} = \left\langle \psi_0^{(0)} \middle| H^{(0)} + \delta H \middle| \psi_0^{(0)} \right\rangle = f_H\left(\psi_0^{(0)}\right) , \tag{9.29}$$

and since $\psi_0^{(0)} \in \{\psi_\lambda\}$, this is an upper bound for the infimum of f_H on $\{\psi_\lambda\}$. ■

9.4 Approximating excited states

Though the ground state energy tends to be of particular interest, we may also want to approximate the energies of excited states. If we were to have perfect knowledge of the first k eigenstates $\psi_0, \ldots \psi_{k-1}$, with energies $E_0, E_1, \ldots, E_{k-1}$, say, then it would be straightforward to use the same variational ideas as we did above to approximate the $k+1$st state. Letting $\mathcal{K} = \text{Span}\{\psi_0, \ldots, \psi_{k-1}\}$, we would have the following.

Proposition 9.4.1. If $\inf_{\mathcal{K}^\perp} f_H$ is attained on for some $\psi_k \in \mathcal{K}^\perp$, then this is the $k+1$st lowest energy eigenstate, and $f_H(\psi_k) \geqslant E_j, j = 0, \ldots, k-1$.

Proof. Since for our choice of subspace, $H : \mathcal{K} \to \mathcal{K}$, self adjointness implies that $H : \mathcal{K}^\perp \to \mathcal{K}^\perp$. Applying the main variational result for f_H on $\mathcal{K}^\perp$ gives that if f_H achieves its infimum at ψ_k, then $(H - f_H(\psi_k))\left|\psi_k\right\rangle = 0$ and ψ_k is the eigenstate with the lowest eigenvalue in $\mathcal{K}^\perp$. Since the lowest k eigenvalues lie in $\mathcal{K}$ by assumption, this must be the $k+1$st lowest eigenvalue. ■

If this was the end of the story, we would be in the unfortunate situation of needing to have an exact result for lower lying states before approximating the higher ones. It turns out that one can proceed without knowledge of the first k eigenvectors/eigenvalues using the following important theorem.

Theorem 9.4.2 (Min-max)**.** If the infimum,

$$\inf_{\substack{\mathcal{K}\subseteq\mathcal{H}\\ \dim\mathcal{K}=k}}\ \max_{\psi\in\mathcal{K}} f_H(\psi)\ , \tag{9.30}$$

is attained, then it is equal to the kth lowest energy eigenvalue E_{k-1}, and the state vector on which it is attained is a corresponding eigenvector.

We will give the proof for the case where $\mathcal{H}$ has a basis of normalisable energy eigenstates. In the general case (involving generalised eigenstates/continuous spectrum operators), the statement of the theorem and the proof become slightly more technical.

Proof. Let $E_0 \leqslant E_1 \leqslant E_2 \leqslant \ldots$ be the ordered energy eigenvalues and let $\psi_0, \psi_1, \ldots$ be the corresponding stationary state vectors. Let $\mathcal{H}_{k-1} = \mathrm{Span}\{\psi_0, \ldots, \psi_{k-2}\}$. Now for any k-dimensional subspace $\mathcal{K} \subseteq \mathcal{H}$, we will have $\dim\left(\mathcal{K}\cap\mathcal{H}_{k-1}^{\perp}\right) \geqslant 1$ simply for dimensional reasons. Thus, there will be a (normalised) state vector $\psi \in \mathcal{K}$ that can be expressed in terms of the ψ_n for $n \geqslant k-1$,

$$\psi = \sum_{n=k-1}^{\dim\mathcal{H}} a_n\psi_n\ , \qquad \sum_{n=k-1}^{\dim\mathcal{H}} |a_n|^2 = 1\ . \tag{9.31}$$

For this state, we have

$$f_H(\psi) = \sum_{n=k-1}^{d} |a_n|^2 E_n \geqslant \sum_{n=k-1}^{d} |a_n|^2 E_{k-1} = E_{k-1}\ , \tag{9.32}$$

so $\max_{\psi\in\mathcal{K}} f_H(\psi) \geqslant E_{k-1}$.

On the other hand, for $\mathcal{K} = \mathrm{span}\{\psi_0, \psi_1, \ldots, \psi_{k-1}\}$ we have exactly that $\max_{\psi\in\mathcal{K}} f_H(\psi) = E_{k-1}$, so the result follows. ■

By the min-max theorem, the maximum value of f_H on any k-dimensional $\mathcal{K} \subseteq \mathcal{H}$ is an upper bound for the kth energy level of the system. In practice, rather than focus on just the kth energy level, more often than not it makes sense to consider the first k energies collectively in what is sometimes called the *Rayleigh–Ritz method.*

9.5 The Rayleigh–Ritz method

For a given choice of k-dimensional $\mathcal{K} \subseteq \mathcal{H}$, by maximising as in the min-max theorem we can bound the kth energy level. In turn, that maximum value of f_H on $\mathcal{K}$ will be the largest eigenvalue of the restricted Hamiltonian $\Pi_{\mathcal{K}}H\Pi_{\mathcal{K}}$ (this follows immediately from Lemma 9.1.2).

What if, while still restricting to $\mathcal{K}$, we wish to bound the $k-1$st energy level? By the min-max theorem, we will get such a bound if we restrict to any $(k-1)$-dimensional subspace of $\mathcal{K}' \subset \mathcal{K}$ and again maximise the Rayleigh quotient over $\mathcal{K}'$.

What is the optimal choice of such a $\mathcal{K}'$? By an argument identical to that appearing in the proof of the min-max theorem, it is the orthocomplement of the eigenspace of $\Pi_{\mathcal{K}}H\Pi_{\mathcal{K}}$ with largest eigenvalue, and the strongest upper bound we get on the $(k-1)$st energy level ((with $\mathcal{K}$ fixed) is precisely the *second largest eigenvalue of the restricted Hamiltonian* $\Pi_{\mathcal{K}}H\Pi_{\mathcal{K}}$. Continuing the argument *mutatis mutandis*, we arrive at the following procedure.

Definition 9.5.1 (Rayleigh–Ritz variational method (RRVM))**.** Consider a (self-adjoint) Hamiltonian operator H acting on a Hilbert space $\mathcal{H}$ admitting a family of normalisable energy eigenstates with energies bounded below,

$$E_0 \leqslant E_1 \leqslant \cdots E_k \leqslant \cdots\ . \tag{9.33}$$

In the RRVM, we choose a k-dimensional subspace $\mathcal{K} \subseteq \mathcal{H}$ (the *trial subspace* and compute the eigenvalues of the restricted Hamiltonian $\Pi_{\mathcal{K}} H \Pi_{\mathcal{K}}$, which we denote by

$$E_0^{\mathcal{K}} \leqslant E_1^{\mathcal{K}} \leqslant \cdots E_k^{\mathcal{K}} \,. \tag{9.34}$$

Then by the variational arguments above, we have

$$E_n^{\mathcal{K}} \geqslant E_n \,, \qquad n = 0, \ldots, k-1 \,. \tag{9.35}$$

Remark 9.5.2. By another (straightforward) generalisation of the argument used to prove the min-max theorem, one can show that by applying the RRVM for any $k+1$-dimensional subspace $\mathcal{K}^{\vee} \supseteq \mathcal{K}$, one can only *improve* on the estimates derived using $\mathcal{K}$, *i.e.*,

$$E_n^{\mathcal{K}} \geqslant E_n^{\mathcal{K}^{\vee}} \geqslant E_n \,, \qquad n = 0, \ldots, k-1 \,. \tag{9.36}$$

You will explore this dependence of the RRVM results on the size of the trial subspace on Problem Sheet 3.

When applying the RRVM, it is not always convenient to pick one's basis for their trial subspace to be orthonormal from the start. (It is a computational matter to orthonormalise, but in practice we may want to delay doing the computations.) For a not-necessarily-orthonormal basis $\{\varphi_1, \ldots, \varphi_k\}$ for the trial subspace $\mathcal{K}$, the eigenvalue condition is equivalent to the solvability of the system of equations

$$\sum_{i=1}^{k} c_i \left\langle \varphi_j \middle| H - \lambda \middle| \varphi_i \right\rangle = 0 \,, \qquad j = 1, \ldots, k \,,$$

for constants $c_1, \ldots, c_k$. This amounts to the matrix degeneracy condition,

$$\det\left(\left\langle \varphi_j \middle| H \middle| \varphi_i \right\rangle - \lambda \left\langle \varphi_j \middle| \varphi_i \right\rangle \right) \,,$$

which is also sometimes called the *secular equation*.

We see an elementary implementation of the RRVM in the following example.

Example 9.5.3 (Angular momentum (rigid rotor) via Rayleigh–Ritz)**.** The variational methods introduced in this section can be used not just for the Hamiltonian of a system, but for any self-adjoint operator with bounded-below spectrum. Indeed, consider the case of a particle moving on the sphere (the *rigid rotor*). The total angular momentum operator in spherical polar coordinates (and atomic units: $\hbar = 1$), has the form

$$L^2 = -\frac{1}{\sin^2\theta}\left(\sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{\partial^2}{\partial\varphi^2} \right) \,. \tag{9.37}$$

We can obtain Rayleigh–Ritz estimates for the first two eigenvalues of L^2 by taking, as a basis for a space of trial functions $\psi_1(\theta,\varphi) = 1$ and $\psi_2(\theta,\varphi) = \cos^2\theta$. We then need to solve the characteristic equation for the projection of the L^2 operator to this trial space,

$$\det\left(\left\langle \psi_j \middle| L^2 \middle| \psi_k \right\rangle - \lambda \left\langle \psi_j \middle| \psi_k \right\rangle \right) = 0 \,. \tag{9.38}$$

Clearly $L^2\psi_1 = 0$, and we calculate that

$$\begin{aligned} L^2\psi_2 &= -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\cos^2\theta\right) = \frac{2}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin^2\theta\cos\theta\right) \ , \\ &= \frac{2}{\sin\theta}\left(2\sin\theta\cos^2\theta - \sin^3\theta\right) = 2\left(2\cos^2\theta - \sin^2\theta\right) \ , \\ &= 2\left(3\cos^2\theta - 1\right) = 2\left(3\psi_2 - \psi_1\right) \ . \end{aligned} \tag{9.39}$$

We also need the following normalisations and inner products,

$$\begin{aligned} \|\psi_1\|^2 &= \int \sin\theta \ \mathrm{d}\theta\,\mathrm{d}\phi = 4\pi \ , \\ \|\psi_2\|^2 &= \int \cos^4\theta\sin\theta \ \mathrm{d}\theta\,\mathrm{d}\phi = 2\pi\left(-\frac{1}{5}\cos^5\theta\right)\bigg|_0^\pi = \frac{4\pi}{5} \ , \\ \langle\psi_1|\psi_2\rangle &= \int \cos^2\theta\sin\theta \ \mathrm{d}\theta\,\mathrm{d}\phi = 2\pi\left(-\frac{1}{3}\cos^3\theta\right)\bigg|_0^\pi = \frac{4\pi}{3} \ . \end{aligned} \tag{9.40}$$

We clearly have $\left\langle\psi_j\middle|L^2\psi_1\right\rangle = 0 = \left\langle\psi_1\middle|L^2\psi_j\right\rangle$, so the only remaining element is

$$\langle\psi_2|L^2\psi_2\rangle = \langle\psi_2|6\psi_2 - 2\psi_1\rangle = \frac{24\pi}{5} - \frac{8\pi}{3} = \frac{32\pi}{15} \ . \tag{9.41}$$

Our eigenvalue equation therefore gives

$$0 = \begin{vmatrix} -4\lambda\pi & -\frac{4\pi\lambda}{3} \\ -\frac{4\pi\lambda}{3} & \frac{32\pi}{15} - \frac{4\pi\lambda}{5} \end{vmatrix} = \frac{62\pi^2}{45}\lambda(\lambda - 6) \ . \tag{9.42}$$

so $\lambda = 0$ or $\lambda = 6$, which correspond to $\ell = 0$ and $\ell = 2$, respectively, in the standard formula $\ell(\ell + 1)$ for the eigenvalues of L^2 (with $\hbar = 1$). The first case gives the eigenvector $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$, that is ψ_1; the second gives $\left(\begin{smallmatrix}-1\\3\end{smallmatrix}\right)$ or $3\psi_2 - \psi_1$. We recognise these as precisely (up to normalisation) the spherical harmonics Y_0^0 and Y_2^0.

While we have ended up landing on exact eigenstates/eigenvalues, we did not get the first two, but rather the first and fifth lowest eigenstates of L^2—we missed the three $\ell = 1$ states. This is, of course, compatible with our eigenvalues being upper bounds for the first two eigenvalues. However, in our trial basis we explicitly chose only φ-independent functions, so we have effecively enfoced by hand that we are only studying states with $m = 0$. In the $m = 0$ sector, we have ended up with the first and *third* eigenstates. Our trial basis is also invariant under the reflection $\theta \leftrightarrow \pi - \theta$, which restricts to even values of ℓ, and in this subspace we have indeed found exactly the first two eigenstates.

This is frequently a useful trick in applying Rayleigh–Ritz; one specialises to a subspace of the Hilbert space with some definite behaviour with respect to other symmetries of the Hamiltonian and thus can avoid worrying about many lower-energy states with different symmetry properties which are orthogonal to that subspace.

A further elaboration of this method can be implemented by considering *families* of subspaces $\mathcal{K}_\lambda$. In this case, we would attempt to compute the eigenvalues of the restricted Hamiltonian as a function of the parameters $\{\lambda\}$, and then the best bound for a given eigenvalue would come from minimising as a function of λ. In this case, the best bound for two different energy levels E_m and E_n ($n \neq m$) might very well come from different values of λ.

Chapter 10

WKB Approximation

We now come to a much different kind of approximation, the *semi-classical approximation* for stationary state wave functions. This is also known as the WKB approximation, in honour of physicists Wentzel, Kramers, and Brillouin who developed the method in the mid 1920's. Unlike the approximation methods of the previous chapters, which were formulated in the abstract language of Hilbert spaces and observables, the semi-classical approximation that we will consider here is very much tailored to the study of wave functions in particular. Indeed, there is an entire branch of the analysis of PDEs known as *semi-classical analysis* that is closely related to the methods presented here.

We begin by describing an informal "derivation" of what we will soon come to understand as the zeroth order WKB approximation. Recall that the momentum operator acts on wave functions according to

$$(P\psi)(x) = -i\hbar\psi'(x)\ . \tag{10.1}$$

Now for a given potential energy function $V(x)$ and a given energy E, the *classical* momentum of a particle with that energy at a given x (assuming $E > V(x)$) would be given by

$$p(x) = \sqrt{2m(E - V(x))}\ . \tag{10.2}$$

One might then imagine that a wave function for a state with energy E would obey something like an equation of the form

$$P\psi(x) \overset{?}{=} \pm p(x)\psi(x)\ , \tag{10.3}$$

which is just a first order ordinary differential equation. This can be solved directly as follows,

$$\psi(x) \overset{?}{=} \exp\left(\pm\frac{i}{\hbar}\int_{x_0}^{x} p(s)\,\mathrm{d}s\right)\ . \tag{10.4}$$

In general, this analysis is obviously flawed; in particular, when we evaluate the kinetic energy operator $P^2/2m$ on such a wave function, the second action of P will not only bring down another copy of $p(x)$ but by the product rule will also differentiate $p(x)$. Consequently, this analysis exactly valid only when $p(x)$ is a constant, in which case we just have a plane wave solution, *i.e.*, a generalised momentum/energy eigenstate.

Nevertheless, there is some appeal to the idea that the operator P should more or less look like the classical momentum as a function of x, at least in some kind of limit. Indeed, if there is a limiting situation in which quantum mechanics starts to systematically reduce to classical mechanics, one might very well expect such a relation to hold. It turns out there is often such a limit—it is known as the *semi-classical limit*—and the above *ad hoc* wave function is just the first approximation in a systematic expansion.

10.1 The semi-classical expansion and WKB approximation

The starting point for making the previous procedure more systematic is to rewrite a stationary state wave function in terms of (the exponential of) its logarithm,

$$\psi(x) = \exp\left(\frac{iS(x)}{\hbar}\right)\ . \tag{10.5}$$

In light of the heuristic discussion before, we anticipate that the phase S might be related to the integral of the classical momentum in some regime. The time-independent Schrödinger equation in terms of this polar expression takes the form (after dividing through by $\psi(x)$ and rearranging some terms),[56]

$$S'(x)^2 - i\hbar S''(x) = 2m(E - V(x)) = p^2(x)\ . \tag{10.6}$$

The key assumption that we make at this point is that *as an expansion in* $\hbar$, we have

$$S(x) = S^{(0)}(x) + \hbar S^{(1)}(x) + \dots\ . \tag{10.7}$$

This is sometimes referred to as a *semi-classical expansion*, since the parameter $\hbar$ can be thought of as characterising a scale where quantum effects become important, so the $\hbar \to 0$ limit should in some sense be a classical limit.[57]

Solving order by order in $\hbar$, we find that the first two terms in the semi-classical expansion of (10.6) and are given by

$$\begin{aligned} S_0'(x)^2 &= p(x)^2\ , \qquad (10.8)\\ 2S_0'(x)S_1'(x) &= iS_0''(x)\ . \qquad (10.9)\end{aligned}$$

The first equation (10.8) can be solved to give

$$S_0(x) = \pm \int_{x_0}^{x} p(s)\,\mathrm{d}s\ , \tag{10.10}$$

where, as before, p represents the classical momentum as a function of position (and, implicitly, energy). This reproduces our heuristic result (10.4), as promised. Continuing to the first correction (10.9), we compute

$$iS_1'(x) = -\frac{p'(x)}{2p(x)}\ , \tag{10.11}$$

which we can integrate to find

$$iS_1(x) = -\log\left(\sqrt{p(x)}\right)\ . \tag{10.12}$$

The WKB approximation refers to the situation where we truncate the series at this order, giving us the approximate *WKB wave functions*

$$\psi_\pm(x) = \frac{1}{\sqrt{p(x)}}\exp\left(\pm\frac{i}{\hbar}\int_{x_0}^{x} p(s)\,\mathrm{d}s\right)\ . \tag{10.13}$$

In general, this is a local approximation for the wave function and we need to be careful about what happens in the various regions of space, as we will see in a bit. However, there is a simple example where the analysis to this point is entirely sufficient to proceed.

Example 10.1.1 (WKB for particle in a lumpy box). Consider the case of infinite potential barriers at, say, $x = a$ and $x = b$ with $a < b$, and assume $E > V(x)$ for $x \in (a, b)$, though $V(x)$ may be a nontrivial function. We then have WKB wave functions that, by our previous analysis, take the form

$$\psi_{\mathrm{WKB}}(x) = C_+\psi_+(x) + C_-\psi_-(x)\ , \qquad a \leqslant x \leqslant b\ , \tag{10.14}$$

and we need to impose the boundary conditions $\psi(a) = \psi(b) = 0$. Letting $x_0 = a$ in our expressions (10.13),

---

[56] This equation is an instance of the so-called *Riccati equation* for $S'(x)$.

[57] In a physical context, one must be wary about the notion of taking $\hbar \to 0$, since $\hbar$ is a dimensionful parameter with units of angular momentum; one should instead take an appropriate collection of other dimensionful parameters in the problem and form a dimensionless combination involving $\hbar$ that can then be taken to zero by scaling the other variables relative to $\hbar$. For our analysis here it won't be important to keep track of this issue and we will instead treat $\hbar$ as a small parameter; this is what is usuall done in the mathematical treatment of this subject.

the boundary condition at $x = a$ requires that we set $C_+ + C_- = 0$, so we have

$$\psi_{\text{WKB}}(x) = \frac{C}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_a^x p(s)\,\mathrm{d}s\right) . \tag{10.15}$$

Then the requirement $\psi(b) = 0$ gives the *quantisation condition*,

$$\frac{1}{\hbar}\int_a^b p(x)\,\mathrm{d}x = n\pi , \qquad n = 1, 2, 3, \dots . \tag{10.16}$$

For the case of constant potential $V = V_0$, this is just the conventional particle in a box and the WKB wave functions are the true stationary states; (10.16) gives exactly the correct energy levels:

$$\sqrt{2m(E - V_0)}(b - a) = n\pi\hbar \implies E = V_0 + \frac{n^2\pi^2\hbar^2}{2m(b - a)^2} . \tag{10.17}$$

In the case of a non-constant potential, (10.16) gives an approximation to the energy levels of the system.

Estimating the accuracy of the WKB approximation can require some subtle analysis, but to produce a rough proxy for the domain of the validity of the approximation we can inspect when the typical term in the leading equation (10.8) is much larger than the typical term in the subleading equation (10.9),

$$\left(S_0'(x)\right)^2 \gg \hbar \left|S_0''(x)\right| . \tag{10.18}$$

Putting in our solution for $S_0(x)$, we have

$$p(x)^2 \gg \hbar \left|p'(x)\right| , \tag{10.19}$$

which we rewrite presciently as

$$\frac{\hbar}{p(x)^2}\left|p'(x)\right| = \left|\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\hbar}{p(x)}\right)\right| \ll 1 , \tag{10.20}$$

Now if we introduce the *local de Broglie wavelength* $\lambda(x) = h/p = 2\pi\hbar/p$, which represents the wavelength of the generalised momentum eigenstate of momentum $p(x)$, then we have for our condition (dropping a factor of 2π since we are dealing with an extreme inequality),

$$\left|\frac{\mathrm{d}}{\mathrm{d}x}\left(\lambda(x)\right)\right| \ll 1 . \tag{10.21}$$

To reach an intuitive interpretation of our condition, we further multiply again by the de Broglie wavelength,

$$\lambda(x)\left|\frac{\mathrm{d}}{\mathrm{d}x}\left(\lambda(x)\right)\right| \ll \lambda(x) . \tag{10.22}$$

This says that the WKB approximation has a chance of being reliable when the *change of the local de Broglie wavelength over the course of one such wavelength is small compared to that wavelength.* So in terms of percentages, the local de Broglie wavelength should be slowly varying.

We can re-express this condition directly in terms of energies. Using the expression for the classical momentum, we have

$$p'(x) = \frac{mV'(x)}{p(x)} = \frac{m\lambda(x)V'(x)}{2\pi\hbar} , \tag{10.23}$$

which, when we plug it into (10.19), yields the consistency condition (this time dropping a factor of 4π),

$$\left|\lambda(x)V'(x)\right| \ll \frac{p(x)^2}{2m} . \tag{10.24}$$

This says that over the course of a de Broglie wavelength, the potential energy should be slowly varying relative to the

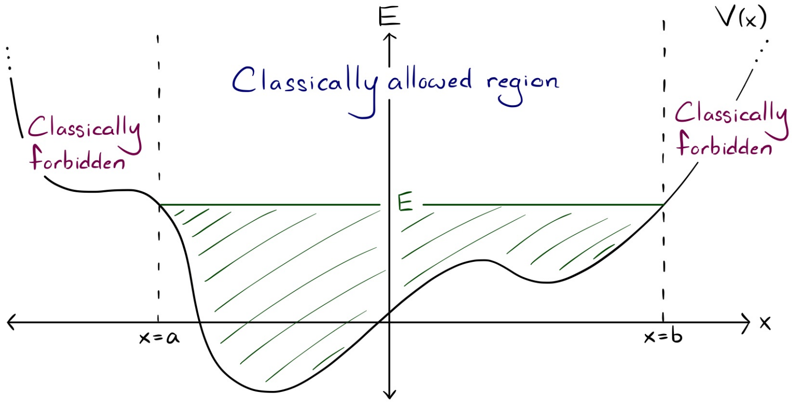


Figure 4. Example of a general potential with a single classically allowed region for the given value of energy (between $x = a$ and $x = b$). A WKB wave function for this type of potential and energy level will be defined in three separate regions and subjected to connection conditions at the classical turning points $x = a$ and $x = b$.

kinetic energy. Thus we expect the WKB approximation to do well for high energies and slow-varying potentials.

10.2 WKB in forbidden regions

For more general potentials (see Figure 4, for example), for a fixed value of E, $E - V(x)$ will become negative for some values of x. These regions in space are referred to as the *classically forbidden regions*, and as $p^2/2m = E - V < 0$, for these regions the classical momentum defined by (10.2) becomes pure imaginary. Instead of the momentum we then introduce the real quantity

$$q(x) = \sqrt{2m(V(x) - E)}\,, \tag{10.25}$$

which is an analogue of the classical momentum in the forbidden region. We then solve (10.8) with an *imaginary* $S_0(x)$,

$$S_0(x) = \pm i \int_{x_0}^{x} q(s)\,\mathrm{d}s\,. \tag{10.26}$$

The $O(\hbar)$ term proceeds analogously, and we arrive at the WKB wave functions for classically forbidden regions,

$$\psi_{\pm}^{\text{forbidden}}(x) = \frac{1}{\sqrt{q(x)}} \exp\left(\pm\frac{1}{\hbar}\int_{x_0}^{x} q(s)\,\mathrm{d}s\right)\,. \tag{10.27}$$

Instead of being oscillatory, these are exponentially growing or decaying as a function of x. Though it is less intuitive, the analysis of validity performed above still applies in this case, with the de Broglie wavelength being replaced by the distance over which the exponentially growing/decaying solution increases/decreases by a factor of e.

10.3 WKB connection formulæ

We assume that, as in the figure, $E - V(x) \geqslant 0$ on the interval $[a, b]$ with $b > a$, and is negative outside and vanishes at a and b. These two points are referred to as the *classical turning points*, in reference to the classical trajectory at this energy. In order for our approximate solution to be normalisable, the solution in the left-most classically forbidden region must be exponentially growing with x (so decaying as $x \to -\infty$), and in the right-most forbidden region must be exponentially decaying with x. We therefore seek a solution of the form[58]

$$\psi(x) = \begin{cases} \frac{C_I}{\sqrt{q(x)}} \exp\left(-\frac{1}{\hbar}\int_x^a q(s)\,\mathrm{d}s\right) \ , & x < a \, , \\ \frac{C_+}{\sqrt{p(x)}} \exp\left(\frac{i}{\hbar}\int_a^x p(s)\,\mathrm{d}s\right) + \frac{C_-}{\sqrt{p(x)}} \exp\left(-\frac{i}{\hbar}\int_a^x p(s)\,\mathrm{d}s\right) \ , & a < x < b \, , \\ \frac{\tilde{C}_+}{\sqrt{p(x)}} \exp\left(\frac{i}{\hbar}\int_x^b p(s)\,\mathrm{d}s\right) + \frac{\tilde{C}_-}{\sqrt{p(x)}} \exp\left(-\frac{i}{\hbar}\int_x^b p(s)\,\mathrm{d}s\right) \ , & a < x < b \, , \\ \frac{C_{II}}{\sqrt{q(x)}} \exp\left(-\frac{1}{\hbar}\int_b^x q(s)\,\mathrm{d}s\right) \ , & x > b \, . \end{cases} \tag{10.28}$$

There is a key subtlety having to do with deciding how to *connect* the exponentially increasing/decreasing solutions across the classical turning points at a and b to the oscillatory WKB wave functions in the classically allowed region. Indeed, all of our WKB wave functions actually diverge at a and b due to the denominator vanishing when $E = V(x)$. This represents is a *breakdown in the WKB approximation* in the vicinity of classical turning points.

To investigate the situation, we perform an additional approximate analysis in a *small neighbourhood of the classical turning point*. For $x \approx b$, say, we approximate the potential (assuming it is sufficiently smooth) as a linear function,

$$V(x) \approx V(b) + (x-b)V'(b) \, , \tag{10.29}$$

where in this case $V'(b)$ is positive. We then consider the Schrödinger equation for this approximation. Setting $y = x - b$ and $\tilde{\psi}(y) = \psi(x)$, we have

$$-\frac{\hbar^2}{2m}\tilde{\psi}''(y) = (E - V(b) - yV'(b))\,\tilde{\psi}(y) = -yV'(b)\tilde{\psi}(y) \, . \tag{10.30}$$

Introducing a further variable $z = \left(2mV'(b)/\hbar^2\right)^{1/3} y$ and defining $\varphi(z) = \tilde{\psi}(y)$, this becomes a famous ordinary differential equation, the *Airy equation*,

$$\varphi''(z) = z\varphi(z) \, . \tag{10.31}$$

We will take for granted the following integral expressions for a basis of solutions of the Airy equation (you can try to confirm for yourself that these solve the Airy equation by differentiating under the integral).

$$\begin{aligned} \mathrm{Ai}(z) &= \frac{1}{\pi}\int_0^\infty \cos\left(\frac{t^3}{3} + zt\right)\,\mathrm{d}t \, , \\ \mathrm{Bi}(z) &= \frac{1}{\pi}\int_0^\infty \left(\sin\left(\frac{t^3}{3} + zt\right) + \exp\left(-\frac{t^3}{3} + zt\right)\right)\,\mathrm{d}t \, . \end{aligned} \tag{10.32}$$

What will be important for our purposes here is the large $|z|$ asymptotics of these functions, which take the form (again, feel free to take this for granted),

$$\mathrm{Ai}(z) \sim \frac{\exp\left(-\frac{2}{3}z^{\frac{3}{2}}\right)}{2\sqrt{\pi}z^{\frac{1}{4}}} \, , \qquad \mathrm{Bi}(z) \sim \frac{\exp\left(\frac{2}{3}z^{\frac{3}{2}}\right)}{\sqrt{\pi}z^{\frac{1}{4}}} \, , \qquad z \gg 1 \, , \tag{10.33}$$

$$\mathrm{Ai}(z) \sim \frac{\cos\left(\frac{2}{3}(-z)^{\frac{3}{2}} - \frac{\pi}{4}\right)}{\sqrt{\pi}(-z)^{\frac{1}{4}}} \, , \quad \mathrm{Bi}(z) \sim \frac{\cos\left(\frac{2}{3}(-z)^{\frac{3}{2}} + \frac{\pi}{4}\right)}{\sqrt{\pi}(-z)^{\frac{1}{4}}} \, , \quad z \ll -1 \, , \tag{10.34}$$

We see that it is $\mathrm{Ai}(z)$ that behaves like a decaying exponential for large positive z, while $\mathrm{Bi}(z)$ instead behaves like a growing exponential. This suggests that we should want to use the $\mathrm{Ai}(z)$ solution to interpolate between the forbidden and allowed regions.[59]

---

[58] Notice the strategic choice of limits of integration adopted here. We have given two versions of the wave function in the classically allowed region, one adapted for comparing to the left-most forbidden region and the other adapted for comparing to the right-most forbidden region.

[59] In a careful treatment, we should further subdivide our space to include turning point regions where we use this Airy approximation, and these should overlap with the regions where the WKB wave functions are valid. This level of detail is important for an estimation of the size of errors in the WKB approximation, but will not be necessary for us.

Indeed, if we consider the forbidden-region decaying exponential WKB wave function in the right region and use the same approximation (10.29) for the potential near $x = b$, then we find

$$
\begin{aligned}
\frac{C_{II}}{\sqrt{q(x)}} \exp\left(-\frac{1}{\hbar}\int_b^x q(s)\,\mathrm{d}s\right) &\approx \frac{C_{II}}{(2mV'(b)y)^{\frac{1}{4}}} \exp\left(-\left(\frac{2mV'(b)}{\hbar^2}\right)^{\frac{1}{2}}\int_0^y s^{\frac{1}{2}}\,\mathrm{d}s\right) ,\\
&= \frac{C_{II}\exp\left(-\frac{2}{3}z^{\frac{3}{2}}\right)}{(2mV'(b)\hbar)^{\frac{1}{6}}z^{\frac{1}{4}}} ,\\
&\approx C_{II}\,\kappa\,\mathrm{Ai}(z) .
\end{aligned}
\tag{10.35}
$$

where $\kappa = 2\sqrt{\pi}/(2mV'(b)\hbar)^{\frac{1}{6}}$ is a numerical constant that we could also have absorbed into our overall constant. This matches precisely with the asymptotics of the $\mathrm{Ai}(z)$ function up to an overall numerical factor, so we will use $\mathrm{Ai}(z)$ to interpolate from the forbidden to the allowed region.

In the allowed region, then, we have to match to an appropriate combination of oscillatory WKB wave functions. To this end, we observe that in the allowed region we have, under the approximation (10.29) for $x \approx b$,

$$
\begin{aligned}
\frac{2C_{II}}{\sqrt{p(x)}} \cos\left(\frac{1}{\hbar}\int_x^b p(s)\,\mathrm{d}s - \frac{\pi}{4}\right) &\approx \frac{2C_{II}}{(-2mV'(b)y)^{\frac{1}{4}}} \cos\left(\left(\frac{2mV'(b)}{\hbar^2}\right)^{\frac{1}{2}}\int_y^0 (-s)^{\frac{1}{2}}\,\mathrm{d}s - \frac{\pi}{4}\right)\\
&= \frac{2C_{II}\cos\left(\frac{2}{3}(-z)^{\frac{3}{2}} - \frac{\pi}{4}\right)}{(2mV'(b)\hbar)^{\frac{1}{6}}(-z)^{\frac{1}{4}}} ,\\
&\approx \kappa\, C_{II}\mathrm{Ai}(z) .
\end{aligned}
\tag{10.36}
$$

We conclude that to interpolate with the $\mathrm{Ai}(z)$ Airy function, we should choose $\tilde{C}_{\pm}$ in the allowed region so that they combine to give the first expression in (10.36). An analogous treatment at the turning point $x = a$ implies that the allowed-region WKB wave function on the right hand side of that turning point should be given by

$$
\psi(x) = \frac{2C_I}{\sqrt{p(x)}} \cos\left(\frac{1}{\hbar}\int_a^x p(s)\,\mathrm{d}s - \frac{\pi}{4}\right) .
\tag{10.37}
$$

The resulting connection formulæ are summed up in the following.

Proposition 10.3.1. For continuation to the exponentially decreasing solution past the turning point at b we must have

$$
\tilde{C}_+ = C_{II}\,\mathrm{e}^{-\frac{\pi i}{4}} , \quad \tilde{C}_- = C_{II}\,\mathrm{e}^{\frac{\pi i}{4}} \quad \Longrightarrow \quad \psi(x) = \frac{2C_{II}}{\sqrt{p(x)}} \cos\left(\frac{1}{\hbar}\int_x^b p(s)\,\mathrm{d}s - \frac{\pi}{4}\right) ,
\tag{10.38}
$$

Similarly, for continuation to the solution that exponentially decays as $x \to -\infty$ past the turning point at a we must have

$$
C_+ = C_I\,\mathrm{e}^{-\frac{\pi i}{4}} , \quad C_- = C_I\,\mathrm{e}^{\frac{\pi i}{4}} \quad \Longrightarrow \quad \psi(x) = \frac{2C_I}{\sqrt{p(x)}} \cos\left(\frac{1}{\hbar}\int_a^x p(x)\,\mathrm{d}s - \frac{\pi}{4}\right) ,
\tag{10.39}
$$

Remark 10.3.2. Though not important in this particular analysis, one does run into situations where one wants to match onto the exponentially *growing* solution on the other side of the classical turning point. In this case we have, by an analogous analysis, that if the wave functions in the forbidden regions are of the form

$$
\psi_I(x) = \frac{D_I\exp\left(\frac{1}{\hbar}\int_x^a q(s)\,\mathrm{d}s\right)}{\sqrt{q(x)}} , \qquad \psi_{II}(x) = \frac{D_{II}\exp\left(\frac{1}{\hbar}\int_b^x q(s)\,\mathrm{d}s\right)}{\sqrt{q(x)}} ,
\tag{10.40}
$$

then the matching must be done with the Bi(z) Airy function and one has in the classically allowed region

$$\tilde{C}_+ = \frac{D_{II}}{2}\,\mathrm{e}^{\frac{\pi i}{4}}\,, \quad \tilde{C}_- = \frac{D_{II}}{2}\,\mathrm{e}^{-\frac{\pi i}{4}} \quad \Longrightarrow \quad \psi(x) = \frac{D_{II}}{\sqrt{p(x)}}\cos\left(\frac{1}{\hbar}\int_x^b p(s)\,\mathrm{d}s + \frac{\pi}{4}\right)\,, \tag{10.41}$$

for matching to the right and

$$C_+ = \frac{D_I}{2}\,\mathrm{e}^{\frac{\pi i}{4}}\,, \quad C_- = \frac{D_I}{2}\,\mathrm{e}^{-\frac{\pi i}{4}} \quad \Longrightarrow \quad \psi(x) = \frac{D_I}{\sqrt{p(x)}}\cos\left(\frac{1}{\hbar}\int_a^x p(x)\,\mathrm{d}s + \frac{\pi}{4}\right)\,, \tag{10.42}$$

for matching to the left. Taken together, this full set of connection formulæ allow us to match an arbitrary solution across turning points.

Remark 10.3.3. An important feature of this result is that, when all is said and done, we can forget about the Airy functions and the interpolation region; the relation between the coefficients in the allowed and forbidden regions is fixed *universally* subject to only the assumption that the potential is smooth at the turning points.

Remark 10.3.4. There is another way of deducing these connection formulæ that is quite elegant, though the justification is not entirely transparent. The idea is to *analytically continue* the WKB wave functions around the classical turning point, avoiding the singularity, and matching on either side. In other words, for (say) the turning point at $x = a$, set $x - a = \rho e^{i\varphi}$, with ρ sufficiently large that the WKB approximation can plausibly stays reliable. Starting with the exponential solution in the forbidden region, we continue along the path in the upper half plane ($\varphi \in (0, \pi)$) and this produces the coefficient C_- near a; the C_+ term is instead obtained by analytic continuation in the lower half plane ($\varphi \in (\pi, 2\pi)$). In this treatment, the important phase shift by $\pi/4$ arises from the analytic continuation of the $1/\sqrt{p} \simeq (x-a)^{-\frac{1}{4}}$ factor. A similar analysis follows at $x = b$.

10.4 Bohr–Sommerfeld quantisation

We produced two expressions for the WKB wave function in the classically allowed region by matching to the appropriate exponential wave functions in both forbidden regions. The requirement that these two expressions agree gives the *Bohr–Sommerfeld quantisation rule*, which generalises the quantisation condition from our example to the case with finite potential in the classically forbidden regions.

Corollary 10.4.1 (Bohr–Sommerfeld quantisation rule)**.** Normalisable semiclassical solutions satisfying the connection formulæ at classical turning points exist if and only if

$$\int_a^b p(x)\,\mathrm{d}x = \left(n + \frac{1}{2}\right)\pi\hbar\,. \tag{10.43}$$

Proof. Equating the two expressions for the allowed-region WKB wave function we have

$$\frac{C_I}{\sqrt{p}}\cos\left(\frac{1}{\hbar}\int_a^x p(s)\,\mathrm{d}s - \frac{\pi}{4}\right) = \frac{C_{II}}{\sqrt{p}}\cos\left(\frac{1}{\hbar}\int_x^b p\,\mathrm{d}s - \frac{\pi}{4}\right)\,. \tag{10.44}$$

Rewriting the argument of the cosine on the right hand side, we have

$$\frac{1}{\hbar}\int_x^b p(s)\,\mathrm{d}s - \frac{\pi}{4} = \frac{1}{\hbar}\int_a^b p(s)\,\mathrm{d}s - \frac{1}{\hbar}\int_a^x p(s)\,\mathrm{d}s - \frac{\pi}{4}\,. \tag{10.45}$$

Further using the fact that cosine is an even function, we have that one of the following must hold

$$\begin{aligned} C_I &= +C_{II}\,, \qquad \frac{1}{\hbar}\int_a^b p(x)\,\mathrm{d}x = \frac{\pi}{2} + 2n\pi\,, \qquad n = 0,1,2,\ldots\,, \\ C_I &= -C_{II}\,, \qquad \frac{1}{\hbar}\int_a^b p(x)\,\mathrm{d}x = \frac{\pi}{2} + (2n+1)\pi\,, \qquad n = 0,1,2,\ldots\,. \end{aligned} \tag{10.46}$$

Allowing for either sign this gives the expected condition,

$$\int_a^b p(x)\,\mathrm{d}x = \left(n + \frac{1}{2}\right)\pi\hbar\,, \qquad n = 0,1,2,\ldots\,. \tag{10.47}$$

The correction factor of 1/2 coming from the connection conditions is known as the *Maslov correction*. ■

This condition is capable of giving surprisingly good answers. For example, it is *exact* for the simple harmonic oscillator.

A common interpretation/application/perspective on (10.43) arises from expressing the same quantity as an area integral. Indeed, if we identify the region $A(E) \subset \mathbb{R}^2_{x,p}$ where $p^2 \leqslant 2m(E - V(x))$, then we estimate the *number of quantum states* corresponding to the classical states whose trajectories are confined to this region in phase space by

$$\#\,\text{states}(E) \approx n(E) = \frac{1}{\pi\hbar}\int_{a(E)}^{b(E)} p\,\mathrm{d}x = \frac{1}{2\pi\hbar}\iint_{A(E)} \mathrm{d}p\,\mathrm{d}x\,, \tag{10.48}$$

where the final equality involves a factor of two because the area of the region includes both the area above the x-axis and the area below it. Since wave functions decay exponentially fast outside the region, this number can also be thought of as an estimate of the number of states whose wave functions are supported in $A(E)$.

This formula is often summarised by saying that there is, roughly, a quantum state for each $2\pi\hbar$ unit of area in phase space; this can be generalised to systems in higher dimensions, in which case there is roughly one quantum state for each $(2\pi\hbar)^d$ unit of volume in phase space.

10.5 The radial WKB approximation

The WKB method we've been studying is particularly suited to the case of one-dimensional systems. We can easily extend this to three dimensional problems in the case where spherically symmetry allows us to restrict to definite angular momentum eigenstates and then solve a one-dimensional radial problem. Indeed, with central potential $V(\mathbf{x}) = V(r)$, we have for $\psi(\mathbf{x}) = R(r)Y_\ell^m(\theta,\varphi)$ the radial (time-independent) Schrödinger equation,

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}(rR)\right] + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}R(r) = (E - V(r))\,R(r)\,, \tag{10.49}$$

which can be rewritten as a one-dimensional Schrödinger equation for $rR(r)$ (with a modified potential for nonzero angular momentum),

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2}(rR) = \left(E - V(r) - \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\right)(rR)\,. \tag{10.50}$$

Consequently we have radial WKB wave functions given by

$$R_\pm(r) = \frac{1}{rp(r)^{\frac{1}{2}}}\exp\left(\pm\frac{i}{\hbar}\int^r p(r)\right)\,, \tag{10.51}$$

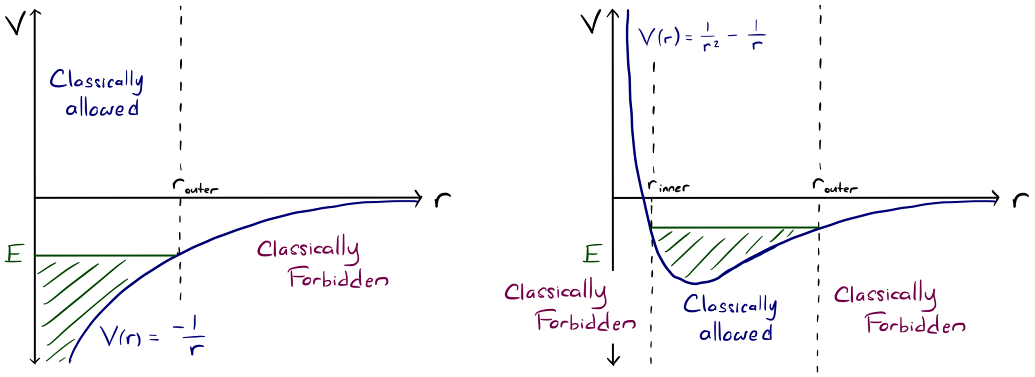


Figure 5. Radial potentials (in this case similar to the Coulomb potential) with and without "centrifugal" term from angular momentum. In the case without, there is a single classical turning point at r_{outer}, while for the case with angular momentum there is also an inner turning point at r_{inner}.

where

$$p(r)^2 = 2m\left(E - V(r) - \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\right) . \tag{10.52}$$

In the case where $V(r)$ is a strictly increasing function of r (such as the harmonic oscillator or the Hydrogen atom), there is an important distinction between the case where $\ell = 0$ (spherically symmetric states), for which there is no inner turning point, and the case where $\ell \neq 0$, for which for any energy there will be an inner turning point as long as the potential diverges less than quadratically with radius at the origin. (See Figure 5.)

The semiclassical wave function must still satisfy the connection conditions of 10.3.1 at $r = r_{\text{outer}}$. However, there is a new ingredient in the case when $\ell = 0$, which is that for $R(r)$ to be bounded, $rR(r)$ should vanish at the origin. As a result, we must have the sin combination of $R_\pm$ wave functions,

$$R(r) = \frac{C}{rp(r)^{\frac{1}{2}}}\sin\left(\frac{1}{\hbar}\int_0^r p(s)\,\mathrm{d}s\right) = \frac{\tilde{C}}{rp(r)^{\frac{1}{2}}}\cos\left(\frac{1}{\hbar}\int_r^{r_{\text{outer}}} p(s)\,\mathrm{d}s - \frac{\pi}{4}\right) , \tag{10.53}$$

and to match both expressions we need

$$\frac{1}{\hbar}\int_0^{r_{\text{outer}}} p(s)\,\mathrm{d}s = \left(n + \frac{3}{4}\right)\pi , \qquad n = 0, 1, 2, \ldots . \tag{10.54}$$

For the Hydrogen atom, this yields good estimates for the energies of s-orbitals, as you will see on Problem Sheet 4.

Remark 10.5.1. For states with nonzero angular momentum, one has an inner turning point so there is a naive quantisation condition of the usual form,

$$\frac{1}{\hbar}\int_{r_{\text{inner}}}^{r_{\text{outer}}} p(s)\,\mathrm{d}s = \left(n + \frac{1}{2}\right)\pi , \qquad n = 0, 1, 2, \ldots . \tag{10.55}$$

There is a subtlety here, because the resulting exponentially decaying WKB wave function in the interior forbidden region won't actually be bounded at $r = 0$ due to the enhanced singularity in the effective potential. There is a curious correction known as the *Langer correction* that can be implemented to improve errors arising from this problem at the origin, and you will encounter this as well on Problem Sheet 4.

10.6 Time-dependent WKB*

The relationship between the WKB approximation and classical physics can be drawn out further by considering the analogous approximation to solutions of the time-*dependent* Schrödinger equation. Below we will freely cite concepts from **B7.1 Classical Mechanics**.

Lemma 10.6.1. Parameterising a quantum-mechanical wave function according to

$$\Psi(\mathbf{x},t) = A(\mathbf{x},t)\exp\left(\frac{i}{\hbar}S(\mathbf{x},t)\right), \tag{10.56}$$

where A and S are both real, the Schrödinger equation with Hamiltonian $H = P^2/2m + V(X)$ is equivalent to the following pair of equations,

$$\begin{aligned} \frac{\partial S}{\partial t} + \frac{|\nabla S|^2}{2m} + V &= \frac{\hbar^2}{2m}\frac{\nabla^2 A}{A}, && (10.57) \\ \frac{\partial A^2}{\partial t} + \nabla\cdot\left(\frac{A^2}{m}\nabla S\right) &= 0. && (10.58) \end{aligned}$$

Proof. Direct calculation yields

$$\nabla\Psi = \left(\frac{\nabla A}{A} + \frac{i}{\hbar}\nabla S\right)\Psi, \qquad \frac{\partial\Psi}{\partial t} = \left(\frac{1}{A}\frac{\partial A}{\partial t} + \frac{i}{\hbar}\frac{\partial S}{\partial t}\right)\Psi, \tag{10.59}$$

and continuing,

$$\nabla^2\Psi = \left(\frac{\nabla^2 A}{A} + \frac{i}{\hbar}\nabla^2 S + 2\frac{i}{\hbar}\frac{\nabla A}{A}\cdot\nabla S - \frac{1}{\hbar^2}|\nabla S|^2\right)\Psi. \tag{10.60}$$

Substituting these into Schrödinger's equation and dividing by Ψ yields a complex equation whose real and imaginary parts are, after a little manipulation, the desired pair of equations. ■

The probability density is $|\Psi|^2 = A^2$ and the probability current is

$$\mathbf{j} := i\frac{\hbar}{2m}\left(\Psi\nabla\overline{\Psi} - \overline{\Psi}\nabla\Psi\right) = \frac{A^2}{m}\nabla S, \tag{10.61}$$

so we can interpret (10.58) as exactly the conservation of probability. The first equation is more subtle to interpret, and is the site of the WKB assumption. Indeed, the $\hbar \to 0$ limit is implemented by ignoring the right hand side of (10.57). This yields:

Definition 10.6.2 (The semi-classical approximation of the time-dependent Schrödinger equation)**.** This determines the wave function $\Psi = Ae^{iS/\hbar}$ satisfying

$$\frac{\partial S}{\partial t} + \frac{|\nabla S|^2}{2m} + V = 0, \tag{10.62}$$

known as the *Hamilton–Jacobi* equation, and

$$\frac{\partial A^2}{\partial t} + \nabla\cdot\left(\frac{A^2}{m}\nabla S\right) = 0, \tag{10.63}$$

the continuity equation.

Remark 10.6.3. This approximation has the best chance to be valid when the right hand side $\hbar^2\nabla^2 A/A$ of (10.57) is small, so in particular, $A \neq 0$.

We recall that the Hamilton–Jacobi equation arises in *classical mechanics* as the equation satisfied by the *action* of the classical trajectory ending at the point $\mathbf{x}$ at time t. For our Hamiltonian, the classical equations of motion are

$$m\ddot{\mathbf{X}} = -\nabla V \, , \tag{10.64}$$

which arise as the Euler–Lagrange equations that follow from extremising the action functional,

$$S[\mathbf{X}(t)] = \int_{t_0}^{t} \mathcal{L}\left(\mathbf{X}(s), \dot{\mathbf{X}}(s)\right)\, \mathrm{d}s \, , \quad \text{where} \quad \mathcal{L}(\mathbf{X}, \dot{\mathbf{X}}) = \frac{m}{2}|\dot{\mathbf{X}}|^2 - V(\mathbf{X}) \, . \tag{10.65}$$

The solution to the Hamilton–Jacobi equation $S(t, \mathbf{x})$ arises as the value of $S[\mathbf{X}_\mathbf{x}(t)]$ when $\mathbf{X}_\mathbf{x}(s)$ are a family of solutions to the classical equations of motion (10.64) chosen so that $\mathbf{X}_\mathbf{x}(t) = \mathbf{x}$. One might, for example, consider the family of trajectories for which $\mathbf{x}(0) = \mathbf{y}$ for a fixed $\mathbf{y}$.

With the suggested boundary condition, evaluating $S(\mathbf{x}, t)$ requires us to integrate along the classical trajectory that joins $\mathbf{y}$ to $\mathbf{x}$. Thus the initial velocity is chosen so that the classical trajectory arrives at $\mathbf{x}$ at time t. The momentum of the trajectory when it passes through $\mathbf{x}$ at time t is then determined by

$$\mathbf{p} = \nabla S \, . \tag{10.66}$$

For a free classical particle (*i.e.*, $V = 0$) our prescription leads to

$$S = \int_0^t \frac{m}{2}\dot{\mathbf{x}}^2 \, \mathrm{d}t = \frac{m|\mathbf{x}-\mathbf{y}|^2}{2t} \, , \qquad \text{where} \qquad \dot{\mathbf{x}} = (\mathbf{x}-\mathbf{y})/t \, . \tag{10.67}$$

It can be easily verified that this satisfies (10.62) with $V = 0$.

Theorem 10.6.4. Given a solution $S(\mathbf{x}, \mathbf{y})$ to the Hamilton–Jacobi equation with the aforementioned boundary conditions, a solution to the continuity equation is given by

$$A^2 = \det\left(\frac{\partial^2 S}{\partial x_j \partial y_k}\right) \, . \tag{10.68}$$

Proof. For simplicity we only prove the one-dimensional case, which follows from direct calculation. The calculation in higher dimensions is more involved but not in a deep way.

$$\begin{aligned} \frac{\partial A^2}{\partial t} &= \frac{\partial^3 S}{\partial x \partial y \partial t} \, , \\ &= -\frac{\partial^2}{\partial x \partial y}\left[\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + V\right] \, , \\ &= -\frac{\partial}{\partial x}\left(\frac{1}{m}\frac{\partial S}{\partial x}\frac{\partial^2 S}{\partial x \partial y}\right) \, , \\ &= -\frac{1}{m}\frac{\partial}{\partial x}\left(\frac{\partial S}{\partial x}A^2\right) \, . \end{aligned}$$

which gives the continuity condition. ■

Thus for the free particle, we obtain

$$A^2 = \det\left(-\frac{m}{t}\mathbb{1}_{3\times 3}\right) = -\left(\frac{m}{t}\right)^3 \, , \tag{10.69}$$

so that the WKB wave function at a future time t is given by

$$\Psi(\mathbf{x}, t) = \left(\frac{m}{t}\right)^{3/2} \exp\left(\frac{im|\mathbf{x}-\mathbf{y}|^2}{2\hbar t}\right) \, . \tag{10.70}$$

Comparing to our result for the propagator (2.53), we see that up to overall normalisation this is exactly the evolution to time t of the generalised position eigenstate for position $\mathbf{y}$ at time zero. That we get the exact answer is actually no surprise, as in this case $\nabla^2 A = 0$, so the WKB equations reproduce the full time-dependent Schrödinger equation.

For time-independent systems and states of definite energy, one can separate out the time dependence and, in the one-dimensional case, recover our previous time-independent analysis.

Proposition 10.6.5. For a time-independent potential $V(\mathbf{x}, t) = V(\mathbf{x})$, the Hamilton–Jacobi equation has solutions of the form

$$S(\mathbf{x}, t) = W(\mathbf{x}) - Et\,, \tag{10.71}$$

provided that

$$\frac{|\nabla W|^2}{2m} + V = E\,. \tag{10.72}$$

The corresponding wave functions,

$$\Psi(\mathbf{x}, t) = A(\mathbf{x}, t) \exp\left(\frac{i(W(\mathbf{x}) - Et)}{\hbar}\right)\,, \tag{10.73}$$

give the approximate eigenstates of energy with eigenvalue E.

Proof. This follows by direct substitution. ■

Chapter 11

One Dimensional Scattering

We have spent most of our time in this course discussing normalisable stationary states, their properties, and the methods used to calculate them. In the application of quantum theory to the real world, there is another large and important subject that has something of a different flavour: the theory of *scattering*.

The basic formulation of the problem is as follows: we imagine that there is some "stuff" that is localised in space, and we want to predict what will happen if we throw some probe particle at the stuff. (Alternatively, one might be interested in observing the result of such throwing-a-particle-at-stuff experiments and reconstructing a microscopic model of the stuff. This is referred to as an *inverse scattering problem*.) The scattering problem arises in both classical dynamics and in quantum mechanics, but of course here we consider the quantum version. In this case one wants to assess the *probability amplitude* for various configurations of outgoing scattered particles.

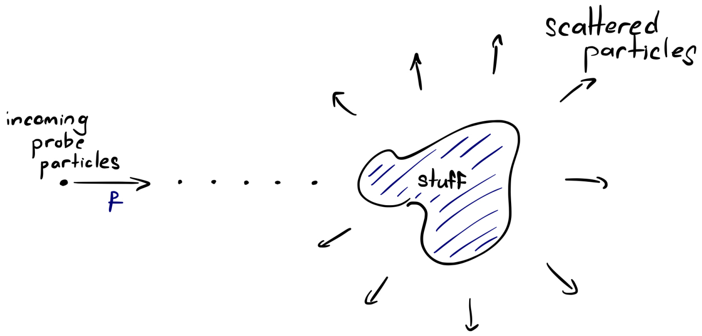


Figure 6. Cartoon representation of a scattering problem.

In the classical setting, we would specify the asymptotic trajectory (say, momentum and impact parameter) of the incoming probe particle in the far past (as you will recall from your study of hyperbolic orbits in the Kepler problem in prelims Dynamics) and predict the subsequent trajectory and, in particular, the late-time trajectory when the particle escapes back to infinity.

In the quantum mechanical setting, there is some subtlety in how we realise this intuitive scattering question within our mathematical formalism. The general treatment is quite technical. In this chapter we consider a simplified version of the story, where space is one-dimensional.

11.1 Left-right asymmetric scattering

We consider a situation as depicted in Figure 7, where the potential takes constant values outside of a bounded *interaction region*. The idea is then that particles will propagate freely in the L and R regions, so we can consider particles incident from (say) the left and ask for the amplitude for them to be either reflected back to the left or transmitted through the interaction region out to the right.

To really model the process described above, we would need to perform a time-dependent analysis in which our initial state is a kind of a wave packet localised in the L region and moving to the right, and then we would ask for the late time behaviour of that state. This would require a more involved investigation than we want to pursue for now. Fortunately, it turns out that we can treat this as a time *independent* problem. We consider (generalised) energy eigenstates with

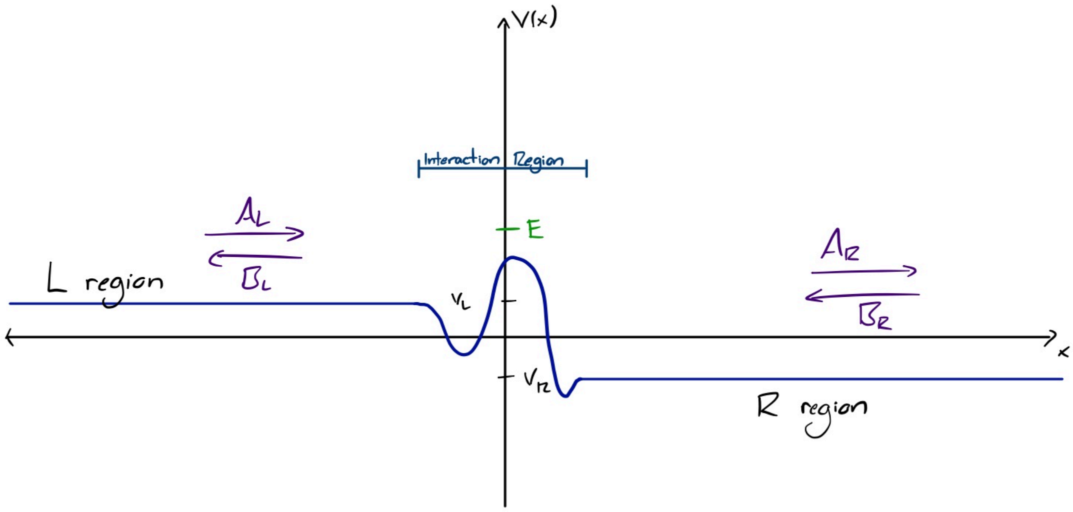


Figure 7. One-dimensional scattering with a localised interaction region.

energy $E > V_{L,R}$, which will necessarily look like plane waves in the L and R regions,

$$\begin{aligned} &\text{for } x \in L\,, \quad \psi(x) = \psi_L(x) = A_L e^{ik_L x} + B_L e^{-ik_L x}\,, \qquad \hbar k_L = p_L\,, \qquad \frac{p_L^2}{2m} = E - V_L\,, \\ &\text{for } x \in R\,, \quad \psi(x) = \psi_R(x) = A_R e^{ik_R x} + B_R e^{-ik_R x}\,, \qquad \hbar k_R = p_R\,, \qquad \frac{p_R^2}{2m} = E - V_R\,. \end{aligned} \tag{11.1}$$

The $A_{L,R}$ terms correspond to the particle having positive momentum, while the $B_{L,R}$ terms describe negative momentum. In the "interaction region" the potential is nontrivial, and it may be difficult to produce an exact expression for the stationary state wave functions there, but on general grounds as the solutions to the time-independent Schrödinger equation we know there will be a two-dimensional space of such wave functions at fixed energy that will interpolate between the plane wave behaviour to the left and the right. Matching onto the solutions in the L and R regions, the detailed form of these solutions will give rise to a linear relationship between the coefficients (A_L, B_L) and (A_R, B_R), which we encode in a (energy-dependent) matrix M,

$$\begin{pmatrix} A_L \\ B_L \end{pmatrix} = M \begin{pmatrix} A_R \\ B_R \end{pmatrix}\,. \tag{11.2}$$

Here we will focus on the case of *scattering from the left*, which we encode by setting $B_R = 0$, as a nonzero value for B_R would be interpreted as indicating some nonzero probability for the particle to be *arriving* from the right. In this case, we define the following physically important quantities.

Definition 11.1.1. The *reflection coefficient* R and the *transmission coefficient* T are defined (as functions of energy) for one-dimensional scattering according to

$$R = \frac{|B_L|^2}{|A_L|^2}\,, \qquad T = \frac{k_R |A_R|^2}{k_L |A_L|^2}\,. \tag{11.3}$$

These coefficients obey an important conservation condition related to their probabilistic interpretation.

Proposition 11.1.2. The reflection and transmission coefficients are related according to

$$R + T = 1\,. \tag{11.4}$$

We interpret R as the probability that a particle incident from the left with energy E will be reflected off of the potential, and T to be the probability that the particle is transmitted through the potential.

Proof. The simple relation follows from the probability conservation condition for stationary states, which in one dimension reads as§

$$\partial_x j(x) = 0 \ , \qquad j(x) = \frac{\hbar}{2mi}\left(\overline{\psi(x)}\partial_x\psi(x) - \psi(x)\partial_x\overline{\psi}(x)\right) \ . \tag{11.5}$$

You have encountered this conservation rule in your **A11 Quantum Theory**, and it follows as an immediate consequence of the time-independent Schrödinger equation. Applying this condition to stationary *scattering states* as above, we have

$$j(x) = \begin{cases} \frac{p_L}{m}|A_L|^2 - \frac{p_L}{m}|B_L|^2 \ , & x \in L \ , \\ \frac{p_R}{m}|A_R|^2 - \frac{p_R}{m}|B_R|^2 \ , & x \in R \ . \end{cases} \tag{11.6}$$

Conservation of the probability current then equates the value of $j(x)$ on either side of the interaction region and gives

$$\frac{p_L}{m}|A_L|^2 + \frac{p_R}{m}|B_R|^2 = \frac{p_L}{m}|B_L|^2 + \frac{p_R}{m}|A_R|^2 \ . \tag{11.7}$$

setting $B_R = 0$ and dividing through by the left hand side gives $R + T = 1$. ■

Remark 11.1.3. Equation (11.7) is often understood in slightly different terms by making a somewhat different (and non-canonical) interpretation of these generalised energy eigenstates. If we say that a wave function of the form

$$\psi(x) = Ae^{\frac{ipx}{\hbar}} \ , \tag{11.8}$$

describes an *ensemble of particles* (sometimes people say a *beam of particles*) travelling with momentum p and density $|A|^2$, then the flow rate of these particles will be given by $\frac{p}{m}|A|^2$. In these terms, our probability current j is reinterpreted as an actual flow rate of particles, and the conservation rule becomes a conservation condition for the number of particles in a given region in a steady state: the rate of particles entering into the interaction region (left hand side) is equal to the rate of particles exiting (right hand side).

Remark 11.1.4. It is a remarkable (and not all that obvious) fact that the time-independent analysis given here is sufficient to make predictions about what happens in a more physical scattering setup when one starts with a wave packet approaching the interaction region from the left. The idea is that one can decompose a wave packet in, say, the L region in terms of the scattering states (rather than the usual plane waves of Fourier analysis), and then the time evolution of the wave packet will proceed analogously to what we saw in our discussion of the propagator in Chapter 2. Because the scattering states know about the structure of the interaction region, as the wave packet evolves the it will arrive from the left at the interaction region, do something in the interaction region, and ultimately there will be a reflected and a transmitted wave packet emitted to the left and right, respectively. Importantly, the relative amplitudes will be controlled by R and T (up to the issue of there being a spread of energies in the wave packet, but if the experiment is repeated many times then the law of large numbers dictates that R and T will control the average behaviour, which justifies the "ensemble of particles" interpretation to some extent). A careful analysis of this story goes well beyond our treatment here, but the important conclusion is that this time-independent analysis captures the real physics of the situation!

11.2 Local potential scattering and the S matrix

To have a one-dimensional analogue of higher-dimensional scattering off of a localised potential, it is natural to impose that $V_L = V_R$. (In higher dimensions, if the potential is localised in one region then you can go *around* the potential and so the asymptotic value of the potential should be the same in every direction.) In this case, the conservation condition takes the even nicer form

$$|A_L|^2 + |B_R|^2 = |B_L|^2 + |A_R|^2 \ . \tag{11.9}$$

From a physical point of view (rather than that of solving ODEs), we should be inclined to think of the problem as being that of determining B_R and A_L (the amplitudes of the outgoing parts of the wave function) given A_L and B_R (the amplitudes of the incident parts of the wave function). As long as the upper left-hand component M_{11} of the matrix M,

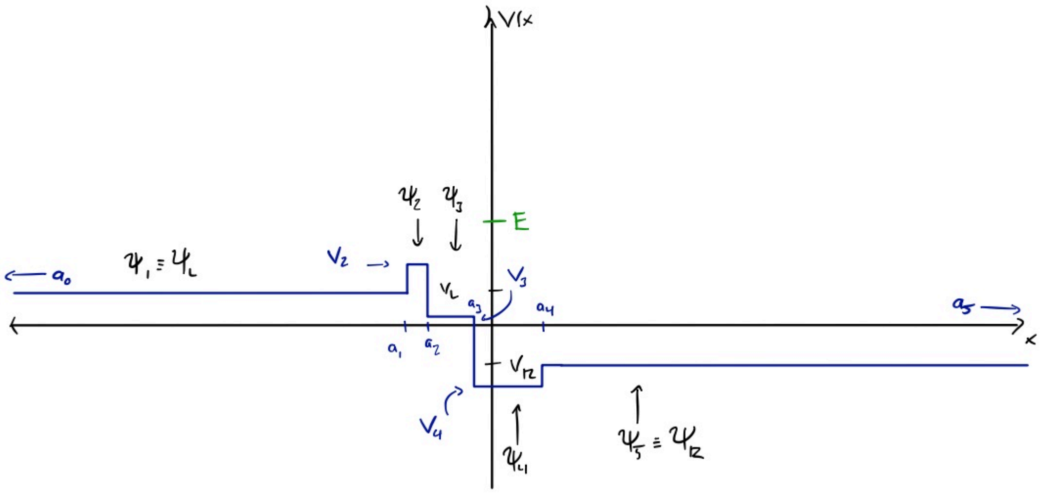


Figure 8. Scattering off of a piecewise constant potential.

we can find such a relation,

$$\begin{pmatrix} A_R \\ B_L \end{pmatrix} = S \begin{pmatrix} A_L \\ B_R \end{pmatrix} , \qquad S = \begin{pmatrix} \frac{1}{M_{11}} & -\frac{M_{12}}{M_{11}} \\ \frac{M_{21}}{M_{11}} & \frac{\det M}{M_{11}} \end{pmatrix} . \tag{11.10}$$

By virtue of (11.9), the matrix S is a norm-preserving endomorphism of $\mathbb{C}^2$ and so a *unitary* 2×2 matrix. Indeed, this is a baby version an important object, the *unitary S-matrix*, which encodes the relationship between incoming and outgoing scattering wavefunctions. (This is an object of significant importance in relativistic quantum field theory and high energy particle physics, where scattering experiments are the main tool of the trade.)

We can then recognise the R and T coefficients in terms of the S matrix coefficients,

$$T = |S_{11}|^2 , \qquad R = |S_{21}|^2 , \tag{11.11}$$

and the condition $R + T = 1$ is a simple consequence of unitarity of S.

Remark 11.2.1. We specialised to scattering from the left, but we could also consider scattering from the right, in which case $A_L = 0$. Then the corresponding reflection and transmission coefficients would be given by $T_{\text{right}} = |S_{22}|^2$ and $R_{\text{right}} = |S_{12}|^2$, which obey an analogous conservation condition.

11.3 Piecewise constant potentials

A (somewhat contrived) class of examples that can be solved exactly, and consequently form a nice test environment for our methods, are the *piecewise constant* potentials (see Figure 8). For these we have a set of junction points $-\infty = a_0 < a_1 < \cdots < a_{n-1} < a_n = \infty$ and set

$$V(x) = V_i , \quad x \in (a_{i-1}, a_i) , \tag{11.12}$$

where in these conventions we have $V_L = V_0$ and $V_R = V_n$. Then our wave function will be piecewise a linear combination of plane waves or exponentials,

$$\psi(x) = \psi_i(x) = A_j e^{ik_j x} + B_j e^{-ik_j x} , \qquad x \in [a_{j-1}, a_j] . \tag{11.13}$$

(If in some region we have $E < V_j$, then we will define $k_j = i\mu_j$ with $\mu_j > 0$. Then the plane wave $e^{ik_j x}$ becomes a decaying exponential $e^{-\mu_j x}$ while $e^{-ik_j x}$ becomes a growing exponential $e^{\mu_j x}$.)

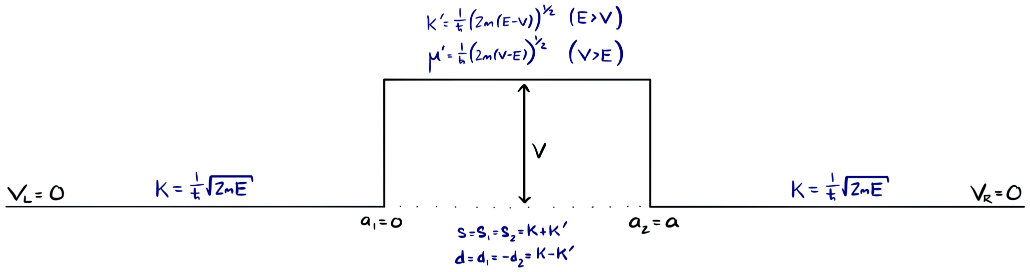


Figure 9. Scattering from a rectangular barrier.

The boundary conditions (continuity of ψ and ψ') at $x = a_i$ require

$$\begin{aligned} \psi_j(a_j) = \psi_{j+1}(a_j) \;&\Rightarrow\; A_j \mathrm{e}^{ik_j a_j} + B_j \mathrm{e}^{-ik_j a_j} = A_{j+1}\mathrm{e}^{ik_{j+1}a_j} + B_{j+1}\mathrm{e}^{-ik_{j+1}a_j} \,, \\ \psi_j'(a_j) = \psi_{j+1}'(a_j) \;&\Rightarrow\; k_j\left(A_j \mathrm{e}^{ik_j a_j} - B_j \mathrm{e}^{-ik_j a_j}\right) = k_{j+1}\left(A_{j+1}\mathrm{e}^{ik_{j+1}a_j} - B_{j+1}\mathrm{e}^{-ik_{j+1}a_j}\right) \,, \end{aligned} \tag{11.14}$$

and this condition can be solved to express the coefficients (A_j, B_j) in terms of (A_{j+1}, B_{j+1}). We encode the relation in a matrix M_j:

$$M_j = \frac{1}{2k_j}\begin{pmatrix} s_j \mathrm{e}^{-id_j a_j} & d_j \mathrm{e}^{-is_j a_j} \\ d_j \mathrm{e}^{is_j a_j} & s_j \mathrm{e}^{id_j a_j} \end{pmatrix} \,, \qquad s_j = k_j + k_{j+1} \,, \quad d_j = k_j - k_{j+1} \,. \tag{11.15}$$

We then have for our total scattering process, $M = M_1 M_2 \cdots M_{n-1}$

Example 11.3.1 (Single barrier scattering and tunnelling). The simplest case of a piecewise constant scattering problem is that of scattering off of a rectangular barrier. In this case there are just two junction points, and as in Figure 9, for ease of notation we will set $a_1 = 0$, $a_2 = a$, $V_L = V_R = 0$, $V_1 = V$. For scattering from the left (in which case $B_R = 0$) we can write

$$\begin{pmatrix} A_L \\ B_L \end{pmatrix} = M_1 M_2 \begin{pmatrix} A_R \\ 0 \end{pmatrix} \,, \tag{11.16}$$

So ultimately we are interested in the left-hand column of the M matrix. Now specialising our general expression for the matrices M_j to our case, we have

$$M_1 = \frac{1}{2k}\begin{pmatrix} s & d \\ d & s \end{pmatrix} \,, \quad M_i = \frac{1}{2k'}\begin{pmatrix} se^{ida} & -de^{-isa} \\ -de^{isa} & se^{-ida} \end{pmatrix} \,, \quad s = k + k' \,, \quad d = k - k' \,, \tag{11.17}$$

which when composed gives us

$$M = \frac{1}{s^2 - d^2}\begin{pmatrix} s^2 e^{ida} - d^2 e^{isa} & sd(e^{-ida} - e^{-isa}) \\ sd(e^{ida} - e^{isa}) & s^2 e^{-ida} - d^2 e^{-isa} \end{pmatrix} \,. \tag{11.18}$$

With some massaging we compute the full S matrix, which is given by

$$\begin{aligned} S &= \frac{1}{d^2 \mathrm{e}^{ias} - s^2 \mathrm{e}^{iad}}\begin{pmatrix} d^2 - s^2 & ds(\mathrm{e}^{-iad} - \mathrm{e}^{-ias}) \\ ds(\mathrm{e}^{iad} - \mathrm{e}^{-ias}) & d^2 - s^2 \end{pmatrix} \,, \\ &= \frac{1}{2ikk'\cos(k'a) + (k^2 + k'^2)\sin(k'a)}\begin{pmatrix} 2ikk'\mathrm{e}^{-ika} & (k^2 - k'^2)\sin(ak')\mathrm{e}^{-2ika} \\ (k^2 - k'^2)\sin(ak') & 2ikk'\mathrm{e}^{-ika} \end{pmatrix} \,. \end{aligned} \tag{11.19}$$

From this we extract the reflection and transmission coefficients,

$$\begin{aligned} T &= \frac{4k^2k'^2}{(k^2+k'^2)^2\sin^2(k'a)+4k^2k'^2\cos^2(k'a)}\,, \\ R &= \frac{(k^2-k'^2)^2\sin^2(k'a)}{(k^2+k'^2)^2\sin^2(k'a)+4k^2k'^2\cos^2(k'a)}\,. \end{aligned} \tag{11.20}$$

As sanity checks, we can observe that as $k' \to k$ (so no barrier), $(T,R) \to (1,0)$, and as $k' \to \infty$ (infinite barrier), $(T,R) \to (0,1)$, and also that the unitarity condition $T+R=1$ does indeed hold here.
To treat the case where $E < V$ transparently, we make the replacement $k' = i\mu'$ with $\mu' > 0$. Being careful with signs coming from imaginary arguments in trigonometric functions, we have

$$\begin{aligned} T &= \frac{4k^2\mu'^2}{(k^2-\mu'^2)^2\sinh^2(\mu'a)+4k^2\mu'^2\cosh^2(\mu'a)}\,, \\ R &= \frac{(k^2+\mu'^2)^2\sinh^2(\mu'a)}{(k^2-\mu'^2)^2\sinh^2(\mu'a)+4k^2\mu'^2\cosh^2(\mu'a)}\,. \end{aligned} \tag{11.21}$$

The most striking result here (though it was clear from the setting up of our problem that this would be the case) is that $T \neq 0$ when $E < V$. This is the phenomenon of *quantum tunnelling*, wherein a particle can transmit through a barrier that would classically block it completely; this behaviour have important technological applications, such as in scanning tunnelling microscopes.

Example 11.3.2 (Bound states and poles). A close relative of our previous example is scattering from a rectangular potential *well*, as in Figure 10. In the first instance, we can simply repurpose our S matrix from the previous example, where now we will have $k' > k$, but otherwise everything will be the same as in (11.19).
The novel feature of this example is that in addition to the scattering states we've been studying, there are also bound states with $V < E < 0$; the bound state wave functions will be of the form

$$\psi_{\text{bound}}(x) = \begin{cases} B_L e^{\mu x}\,, & x < 0\,, \\ A_1 e^{ik'x} + B_1 e^{-ik'x}\,, & 0 < x < a\,, \\ A_R e^{-\mu x}\,, & x > a\,, \end{cases} \tag{11.22}$$

We observe that this is a wave function of precisely the type we considered for scattering states but with the replacement $k = i\mu$, $\mu = \sqrt{-2mE}$ just as in the previous example but now for the wavefunctions in the left and right regions.
Now the bound states correspond to solutions with $A_L = B_R = 0$, which by (11.10) requires that the S matrix become singular. Indeed, upon making the replacement $k \to i\mu$ the (now somewhat formal, as there is no scattering) S matrix takes the form

$$S = \frac{1}{2\mu k'\cos(k'a) + (\mu^2-k'^2)\sin(k'a)} \begin{pmatrix} 2\mu k' e^{\mu a} & (\mu^2+k'^2)\sin(ak')e^{2\mu a} \\ (\mu^2+k'^2)\sin(ak') & 2\mu k' e^{\mu a} \end{pmatrix}\,, \tag{11.23}$$

and each term becomes singular precisely when

$$2\mu k'\cos(k'a) + (\mu^2-k'^2)\sin(k'a) = 0\,. \tag{11.24}$$

Looking back to (11.2), the condition to be able to find a solution with $A_L = B_R = 0$ requires precisely that $M_{11} = 0$, and it is the M_{11} denominator in each entry of the S-matrix that is being set to zero by the condition above. For your own entertainment, you may wish to observe that if instead we take $k \to -i\mu$, then the *same bound states* are responsible for the S matrix developing a kernel.

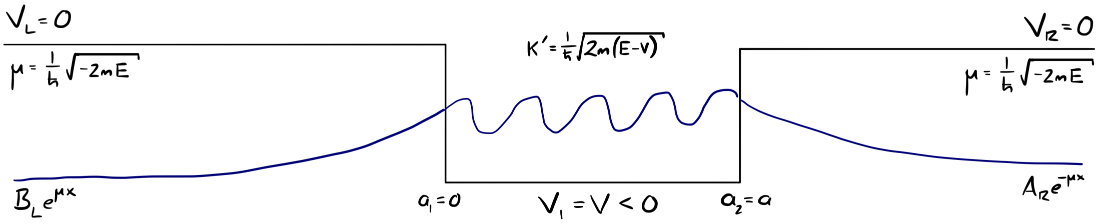


Figure 10. Bound state in a rectangular well.

Remark 11.3.3. What we've observed here is a shadow of a much more general phenomenon in quantum mechanical scattering, where information about bound states can be extracted from the analytic structure (zeroes and poles) of the continuation of scattering data to complex kinematical variables (in this case the asymptotic momentum).

Chapter 12

Epilogue

In this final chapter, we gesture towards some of the topics that, unfortunately, can't be made to fit into an eight week course but nevertheless are equally deserving of discussion. I hope the interested reader will look into them independently.

Time dependent phenomena and methods

In almost the entirety of this course (with the exception of the discussion of the propagator early on) we have aggressively maintained a focus on aspects of the quantum theory that could be studied through time-independent methods. Of course, the world is dynamical and it is often useful to have more intrinsically time-dependent tools at one's disposal. Some key words in these areas are:

- **Pictures of time evolution.** Using the unitary time evolution operator $U(t_1; t_0)$, one can recast the subject of time evolution as applying to the *operators* of quantum theory rather than the states (this is called the *Heisenberg picture* of time evolution). One can go further and evolve states and operators using different time evolution operators (one involving interactions and one corresponding to free propagation). This leads to the *interaction picture*. This formalism is especially important in perturbative scattering theory.

- **Time-dependent perturbation theory.** A realistic, and therefore important, situation to deal with is when a system is perturbed in a manner that is explicitly time-dependent. This could mean that the "underlying" Hamiltonian is time-dependent (say, because you are on the surface of the Earth which is exposed to electromagnetic radiation from the sun periodically), or that we have an underlying time-independent system which we momentarily disturb in a dynamical way (say, by momentarily hitting a Hydrogen atom with a laser beam). This gives rise to slightly different questions than those we addressed in our analysis of perturbation theory. For example, at what rate will the time-dependent perturbation mediate transitions between some given eigenstates of the original system? If you shine a laser at a gas of Hydrogen atoms, how frequently do you expect to ionise the atoms (knock electrons out of bound states into scattering states)? In the case where the time-dependent effect is small, these problems can be treated by a generalisation of perturbation theory to a time-dependent context.

Higher-dimensional scattering

In more than one spatial dimension, the particulars of scattering gets quite a bit more complicated. In particular, the issue of angular dependence takes center stage: given particles incident on a local potential with a fixed momentum, how likely are they to be scattered in any particular direction? This is encoded in something called a differential cross section, and higher-dimensional scattering theory is largely concerned with calculating these cross sections.

The Feynman path integral

An influential "third way" of thinking about quantum theory (in contrast to the algebraic approach of Heisenberg or the differential equation approach of Schrödinger) was supplied by Richard Feynman in a 1948 paper (building on earlier work by himself and others, including Dirac). The idea, roughly, is if we want to compute the propagator,

$$U(x_f, t_f; x_i, t_i) = \left\langle x_f \middle| U(t_f; t_i) \middle| x_i \right\rangle \ . \tag{12.1}$$

then by repeated insertions of resolutions of the identity separated by very short time evolution, one arrives at a picture where one should *sum over all possible trajectories* of the particle between the initial and final position. This sum over histories is encoded in the *Feynman path integral*, which is denoted as follows

$$U(x_f, t_f\,; x_i, t_i) = \int_{x(t_i)=x_i}^{x(t_f)=x_f} [\mathcal{D}x]\mathrm{e}^{\frac{i}{\hbar}S[x(t)]}\ . \tag{12.2}$$

The beautiful result of Feynman's derivation is that the weight with which each trajectory contributes is the (imaginary exponential of) the *classical action* of that trajectory. The integration measure (denoted by $[\mathcal{D}x]$) is a subtle thing to make rigorous sense out of, though in some settings thse subtleties can be overcome. Regardless, the intuition gained from this formulation has proven invaluable for quantum physicists. Indeed, from this perspective, the WKB approximation that we studied in Chapter 9.5 amounts to performing a stationary-phase approximation for the path integral!

You can learn all about path integrals in, for example, **C7.1 Theoretical Physics**.

Entanglement and quantum information theory

We only touched ever-so-briefly upon the issue of quantum entanglement. A more detailed study of the manipulation of finite quantum systems leads to the subject of quantum computing and quantum information theory, in which entanglement is leveraged to perform computational tasks that would seem impossible using conventional classical methods. You can learn all about this in **C7.4 Introduction to Quantum Information**.

Appendix A

Hilbert Space Minutiæ

In the interest of making these notes more self-contained, we provide in this appendix a more in depth account of the definition of and several important properties of Hilbert spaces. The finer points associated with the infinite-dimensional case go beyond the syllabus for the course and are properly treated in a course on Functional Analysis.

A.1 Definitions

Definition A.1.1 (Sesquilinear Form). A *sesquilinear form* on a complex vector space V is a map

$$(\cdot,\cdot) : V \times V \to \mathbb{C}\,,$$

obeying

$$\begin{aligned}(\alpha\varphi_1 + \beta\varphi_2, \psi) &= \bar{\alpha}(\varphi_1, \psi) + \bar{\beta}(\varphi_2, \psi)\,,\\ (\varphi, \alpha\psi_1 + \beta\psi_2) &= \alpha(\varphi, \psi_1) + \beta(\varphi, \psi_2)\,,\end{aligned} \tag{A.1}$$

so it is linear in the second argument and conjugate-linear/$\mathbb{C}$-antilinear in the first argument.[60]

Definition A.1.2 (Hermitian Form). An *Hermitian form* on a complex vector space V is a sesquilinear form that obeys the further *Hermiticity condition*,

$$(\varphi, \psi) = \overline{(\psi, \varphi)}\,. \tag{A.2}$$

This means that for an Hermitian form, $(\psi, \psi) \in \mathbb{R}$ for any $\psi \in V$.

Definition A.1.3 (Hermitian Inner Product). An *Hermitian inner product* on a complex vector space V is a positive definite Hermitian form, *i.e.*, an Hermitian form obeying the additional positive definiteness condition

$$(\varphi, \varphi) \geqslant 0\,, \qquad (\varphi, \varphi) = 0 \iff \varphi = 0\,. \tag{A.3}$$

Such an inner product induces a norm on V,

$$\|\varphi\| := \sqrt{(\varphi, \varphi)}\,. \tag{A.4}$$

Definition A.1.4 (Hilbert Space). A *complex Hilbert space* is a complex vector space equipped with an Hermitian inner product such that all Cauchy sequences converge (so it is *complete*). This means that for any sequence of vectors $\{\varphi_1, \varphi_2, \ldots\}$ such that $\forall\varepsilon > 0$, there exists a natural number $N \in \mathbb{N}$ such that

$$\|\varphi_n - \varphi_m\| < \varepsilon\,, \qquad m, n > N\,, \tag{A.5}$$

there is an element $\varphi \in \mathcal{H}$ such that $\{\varphi_n\} \to \varphi$.

***Definition A.1.5** (Separable Hilbert Space). A *complex Hilbert space* is called *separable* if it admits a countable orthonormal basis.

---

[60] As in the main text, here we adopt “physics conventions” in which the inner product is conjugate-linear in the *first* argument. Often in the mathematics literature the reverse convention is utilised, where the form is conjugate-linear in the second argument and linear in the first.

Every finite-dimensional Hilbert space is separable; separability is a technical condition relevant for the infinite-dimensional case. In this infinite dimensional case, an orthonormal basis is meant in the sense of *infinite* linear combinations,[61] so any vector ψ can be written in terms of the basis vectors $\{\psi_n\}$ as

$$\psi = \sum_{n=1}^{\infty} a_n \psi_n \,, \qquad a_n := (\psi_n, \psi) \,, \tag{A.6}$$

where

$$\sum_{n=1}^{\infty} |a_n|^2 (\psi_n, \psi_n) < \infty \,. \tag{A.7}$$

One is normally only liable to encounter separable Hilbert spaces in quantum mechanical settings, so often the modifier "separable" is omitted entirely.

A.2 Illustrative examples

Example A.2.1 (The Lebesgue space L^2)**.** Wave functions for a particle moving in n dimensions are normally identified with square-integrable functions from $\mathbb{R}^n$ to $\mathbb{C}$, with inner product given by

$$(\psi, \varphi) = \int \overline{\psi(\mathbf{x})} \varphi(\mathbf{x})\, \mathrm{d}^n \mathbf{x} \,. \tag{A.8}$$

Strictly speaking, this space (often denoted $\mathcal{L}^2(\mathbb{R}^n)$) is too big; it includes nonzero functions that are nevertheless zero almost everywhere, and so the Hermitian form doesn't obey the correct positive-definiteness property (there are many functions with zero "norm").

This can be dealt with by taking the quotient by the subspace $\mathcal{N}$ of functions whose norm vanishes, which is equivalent to identifying any two functions that agree almost everywhere. This gives the Lebesgue space,

$$L^2(\mathbb{R}^n) = \left({}^{\mathcal{L}^2(\mathbb{R}^n)} / {}_{\mathcal{N}} \right) \,. \tag{A.9}$$

In this course, we will normally get away with thinking mostly of continuous (or even smooth) wavefunctions and the issue of identifying up to equivalence will not arise.

Example A.2.2 (Incomplete pre-Hilbert space)**.** The requirement of completeness is only relevant for infinite-dimensional Hilbert spaces. To get some intuition for this condition, we can consider the space $C([-1,1],\mathbb{C})$ of continuous, complex-valued functions on the interval $[-1,1]$. This is an Hermitian inner product space with

$$(f, g) = \int_{-1}^{+1} \overline{f(x)} g(x)\, \mathrm{d}x \,. \tag{A.10}$$

This space is incomplete, as can be seen by considering the following family of functions

$$f_n(x) = \begin{cases} 0 & -1 \leqslant x \leqslant -\frac{1}{n} \,, \\ \frac{xn+1}{2} & -\frac{1}{n} \leqslant x \leqslant +\frac{1}{n} \,, \\ 1 & +\frac{1}{n} \leqslant x \leqslant 1 \,. \end{cases} \tag{A.11}$$

This set of functions can be checked to form a Cauchy sequence, but the limit is the discontinuous function that is zero for $x < 0$ and one for $x > 0$. Thus the space of continuous functions with the given inner product is not a Hilbert space; it is sometimes called a *pre-Hilbert space*. The completion of this space, which is the relevant Hilbert space for considering quantum mechanics on the interval, is the Lebesgue space $L^2([-1,1])$ of

---

[61] What is called a *Schauder* basis.

(equivalence classes of) complex-valued, square-integrable functions on $[-1,1]$.

Example A.2.3 (Non-separable Hilbert space)**.** Though they won't show up in this course (or much of anywhere in quantum theory), it may be useful to at least see an example of a Hilbert space that is *not* separable, because seeing is believing. We define the space $\ell^2(\mathbb{R})$ to be the set of functions $f:\mathbb{R}\to\mathbb{C}$ such that $f(x)\neq 0$ for countably many x, and

$$\sum_{x\in\mathbb{R}}|f(x)|^2<\infty\,. \tag{A.12}$$

The inner product is given by

$$(f,g)=\sum_{x\in\mathbb{R}}\overline{f(x)}g(x)\,. \tag{A.13}$$

This admits an uncountable, orthonormal basis which are the functions $\{f_s\}$ with $s\in\mathbb{R}$, where

$$f_s(x)=\begin{cases}1 & x=s\,,\\ 0 & \text{otherwise}\,.\end{cases} \tag{A.14}$$

This clearly has a much different flavour from the sorts of Hilbert spaces we meet when discussing elementary particles or spin systems.

A.3 Operators on Hilbert space

Here we collect some definitions and examples related to the subtleties of operator theory for infinite-dimensional Hilbert spaces. This is only for readers who are independently curious about the subtleties of the rigorous treatment of infinite-dimensional Hilbert spaces; it will not be necessary to concern oneself with this material in the course for the problem sheets or the exam.

Definition A.3.1. An unbounded operator $(A,D(A))$ on a Hilbert space $\mathcal{H}$ is a linear map $A:D(A)\to\mathcal{H}$ from a linear subspace $D(A)\subseteq\mathcal{H}$ to $\mathcal{H}$. It is conventional to require that $D(A)$ is dense in $\mathcal{H}$.

The operators studied in the quantum mechanics of $L^2(\mathbb{R})$ tend to be unbounded operators. For example, the momentum operator P is naturally defined for $D(P)$ the subspace of once-differentiable functions with square-integrable derivatives. The position operator X is likewise defined for functions whose growth at infinity is sufficiently mild that they are still square-integrable after multiplying by x.

Definition A.3.2. The adjoint operator of an unbounded operator $(A,D(A))$ on a Hilbert space $\mathcal{H}$ is another unbounded operator $(A^*,D(A^*))$ on $\mathcal{H}$ obeying

$$(\varphi,A\psi)=(A^*\varphi,\psi)\,,\qquad \forall\psi\in D(A)\,,\ \forall\varphi\in D(A^*)\,. \tag{A.15}$$

The domain $D(A^*)$ is defined to be the linear subspace of $\mathcal{H}$ for which $\varphi\to(\varphi,A\psi)$ is continuous for any $\psi\in D(A)$. By the Riesz–Fréchet isomorphism, $A^*\varphi$ is uniquely defined for a $\varphi\in D(A^*)$ by virtue of the aforementioned continuous linear functional.

In this unbounded setting, a self adjoint operator is, importantly, an operator A for which not only $A=A^*$, but $D(A)=D(A^*)$.

Example A.3.3 (Unbounded operators and adjoints)**.** An instructive example of this subtlety comes in the case of the particle in a box, as reviewed in the prologue of these lecture notes. Recall that here the Hilbert space is $L^2([0,a])$ and the Hamiltonian operator is $H=P^2/2m$.

The Hamiltonian is in fact an unbounded operator, and so needs to be equipped with a choice of domain; the domain used to define the particle in the box is the set of wave functions that vanish at the $x = 0$ and $x = a$ and that are twice differentiable, with the result being square integrable. With this choice of $D(H)$, one finds that $D(H^*) = D(H)$, so the Hamiltonian is truly self-adjoint and the spectral theorem must hold (as you have taken for granted in the past).

On the other hand, consider the momentum operator P for the particle in the box. If we take the same domain $D(P)$ as for H, then (because P is a first order differential operator), $D(P^*)$ is strictly larger than $D(P)$. Indeed, for any $\psi \in D(H)$ and any differentiable φ, we have

$$\int_0^a \overline{\varphi(x)}(P\psi)(x)\,\mathrm{d}x = \int_0^a \overline{P\varphi(x)}(\psi)(x)\,\mathrm{d}x\,, \tag{A.16}$$

so $D(P^*)$ and includes all differentiable functions. Thus we cannot apply the spectral theorem in this case, and indeed we do not have a basis of P-eigenfunctions obeying the prescribed boundary conditions on the interval $[0, a]$.
This example is discussed at length at a quite sophisticated level here.