

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 **All Quantum Theory**. Here we strive for a high degree of accuracy, while stopping short of a completely rigorous discussion (which would require substantial background material from functional analysis). Where we gloss over technical subtleties there will 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 pedagogically illustrate 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 mechanics. 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 (often denoted \mathcal{H}).*

Parsing this postulate requires some definitions, which we reproduce here. ([Appendix A](#) includes for a more detailed review of the definition of Hilbert spaces and some discussion of advanced topics for those who are curious.)

Definition 1.1.1. A complex Hilbert space is a complex vector space \mathcal{H} (possibly infinite-dimensional) with an Hermitian inner product $(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$. Recall that 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,⁹

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

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. A *ray* in \mathcal{H} is the set of non-zero, scalar multiples of a given non-zero vector.

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

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

⁹Observe 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.

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.*,

$$\{\text{Quantum States}\} = (\mathcal{H} \setminus \{0\} / \sim) \cong \mathbb{P}(\mathcal{H}). \quad (1.4)$$

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

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

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} is behind 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 $A : \mathcal{H} \rightarrow \mathcal{H}$ is referred to in the quantum mechanical setting (amongst other place) as an *operator* (as in differential operator). The adjoint of a linear operator $A : \mathcal{H} \rightarrow \mathcal{H}$ is another operator $A^* : \mathcal{H} \rightarrow \mathcal{H}$ such that for any $\varphi, \psi \in \mathcal{H}$ we have,¹⁰

$$(\varphi, A\psi) = (A^*\varphi, \psi). \quad (1.6)$$

There is also an important and, in the general case, difficult 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 quite a bit of technology. *Roughly speaking*, the theorem says that a self-adjoint operator on a Hilbert space always admits a complete (orthonormal) basis of eigenvectors, so for a general observable we can write something like,

$$\psi = \sum_n c_n \psi_n, \quad (1.7)$$

where the ψ_n are eigenstates of the observable A ,

$$A\psi_n = a_n \psi_n. \quad (1.8)$$

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

Remark 1.1.4. The 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 are 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

¹⁰There is an additional complication here in the case of infinite-dimensional \mathcal{H} . General linear operators on an infinite-dimensional \mathcal{H} are only partially defined, so their domain $D(A) \subsetneq \mathcal{H}$ (these are referred to as *unbounded operators*). The adjoint of an operator then has its own domain $D(A^*)$, and self-adjointness requires $D(A) = D(A^*)$ which is not automatic. For our purposes in this course, it will not be important to keep track of domains of the observables we study in any systematic way. See [Appendix A](#) for more.

$\{\psi_{i \in I}\}$ 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

$$\Pi_a \psi = \sum_{i \in I} c_i \psi_i. \quad (1.9)$$

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

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

These probabilistic statements are compatible with the following definitions.

Definition 1.1.5. The *expectation value* of the observable A in a state ψ is given by

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

Definition 1.1.6. The *dispersion* of the observable A in the state ψ is given by

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

These definitions agree with the usual statistical notion of the expectation and variance of a random variable.¹¹

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.

Remark 1.1.7 (Compatible and Incompatible Measurements). If two observables A and B commute, so $[A, B] = 0$,¹² 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,*

$$i\hbar \frac{d\psi}{dt} = H\psi. \quad (1.13)$$

¹¹In the above discussion, our formulæ were all tailored to the case of a normalised state vector ψ . For general ψ , one must divide through by a normalising factor.

¹²Once 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

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,

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

In particular, setting $\varphi = \psi$, this implies that normalised state vectors remain normalised under time evolution.

Let us define an operator $U(t_1; t_0) : \mathcal{H} \rightarrow \mathcal{H}$ that sends a state vector (thought of as the state of our system at time t_0), call it φ_{t_0} , to its time evolution forward to time $t_1 > t_0$, which we call φ_{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*.

Definition 1.1.8. A *unitary operator* U on a Hilbert space is a linear map $U : \mathcal{H} \rightarrow \mathcal{H}$ that obeys

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

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

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

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 The qubit as an instance of the framework

The simplest instances of quantum systems are those with *finite-dimensional Hilbert spaces*. The simplest non-trivial example is then when the Hilbert space has dimension two, in which case the system is often called a *qubit*.¹³ Choosing any orthonormal basis we get a (non-canonical) identification $\mathcal{H} \cong \mathbb{C}^2$ with inner product

$$(\mathbf{u}, \mathbf{v}) = \bar{\mathbf{u}} \cdot \mathbf{v}. \quad (1.21)$$

The space of quantum states in this case is just the complex projective line (*i.e.*, the Riemann sphere), $\mathbb{P}(\mathbb{C}^2) = \mathbb{C}\mathbb{P}^1$. Topologically $\mathbb{C}\mathbb{P}^1 \cong \mathbb{S}^2$, 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 which is as follows,

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

Here σ_0 is just the identity operator and 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 obey the commutator algebra

$$[\sigma_1, \sigma_2] = 2i\sigma_3, \quad [\sigma_2, \sigma_3] = 2i\sigma_1, \quad [\sigma_3, \sigma_1] = 2i\sigma_2. \quad (1.23)$$

In the given basis for \mathcal{H} , σ_3 is diagonalised with eigenvalues ± 1 while σ_1 and σ_2 are not, and the nontrivial commutators imply that these are incompatible observables. By choosing a different basis one could diagonalise σ_1 or σ_2 instead, or indeed 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 $U(2)$ 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 $U(1) \subset U(2)$ of matrices of the form

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

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 $U(2)$ on the space $\mathbb{C}\mathbb{P}^1$ factors through the quotient by this $U(1)$ subgroup, which happens to be isomorphic to the three-dimensional orthogonal group,

$$U(2)/U(1) \cong SO(3). \quad (1.25)$$

There is an obvious action of $SO(3)$ on $\mathbb{C}\mathbb{P}^1$ 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 $SO(3)$ 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 The one-dimensional particle as an instance of the formalism

The primary instance of this formalism treated in **All 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 $\psi : \mathbb{R} \rightarrow \mathbb{C}$, a complex, square-normalisable function of position $x \in \mathbb{R}$. The Hermitian

¹³Qubits play an important role as a building block 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, conventionally called 0 and 1. Unlike classical bits, qubits can live in any linear superposition of their two basis states, which leads to their more powerful computational properties.

inner product of two state vectors φ and ψ is then given by

$$(\varphi, \psi) = \int_{\mathbb{R}} \overline{\varphi(x)} \psi(x) dx. \quad (1.26)$$

Physical wave functions are required to be *normalisable*, so

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

The Hilbert space of such a system is then, roughly speaking, the vector space of complex, square-integrable functions on \mathbb{R}^d . With some technical improvements,¹⁴ 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 we most frequently discuss in this setting are realised as differential operators on wave functions. In particular, the most natural observables are the momentum and position operators P and X , which act according to

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

More generally, we can construct many self-adjoint differential operators as observables by composing the actions of P and X appropriately. (However, recall that because in general $(AB)^* = B^*A^*$, a generic composition of P 's and X 's will not be self-adjoint, though P and X themselves are.) For example, the Hamiltonian in one dimension is usually taken to have its classical form

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

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

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

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

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

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 appearance of the aforementioned subtlety 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 [All Quantum Theory](#) in different terms, are the projection operators

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

where $\mathbf{1}_E(x)$ 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

¹⁴As 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).

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

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

which is exactly how in **All 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 $E \rightarrow \Pi_E$ 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 .