Chapter 1

Quantum mechanics from the ground up

1.1 Introduction

Quantum mechanics is one of our most successful physical theories. It underlies the advances in atomic and solid state physics over the last century, that have propelled forward our technological development, and has provided the backdrop for our explorations of the world of sub-atomic particles and fundamental forces. It is now more than a century since the first components of the theory were identified, and yet it continues to surprise us. Even in the last fifteen years there have been many new discoveries. The development of quantum computation and quantum cryptography have shown that the explicitly non-classical aspects of quantum theory can provide useful and unexpected applications.

This course begins with a revision of the methods of quantum mechanics at a more formal and more general mathematical level than you are likely to have encountered previously. We will then cover some advanced topics in wave-mechanics, such quantum tunnelling in non-idealised potentials, and in time-evolution, from the unitary evolution of isolated systems, through to the non-unitary evolution of quantum systems interacting with their environment.

First, however, we will look again at the postulates of quantum mechanics, and the mathematics we use to write those postulates.

1.2 Postulates of quantum mechanics

When you start a course on special relativity you usually begin by learning two key ideas; Einstein's principle of relativity, that the laws of Physics should be the same in all inertial frames, and the principle of a constant speed of light. From this simple, and physically reasonable starting point, one can derive the laws and equations describing the theory of special relativity. In a similar way, general relativity can be built up from upon some simple physically motivated assumptions.

Quantum mechanics is different. We would like to be able to derive quantum mechanics in a similar way, but we do not (yet) have a set of simple physical principles from which the theory can be derived. Instead, we construct quantum mechanics by writing down a set of mathematical *postulates*. These postulates are not derived from any physical assumption. Instead, as Quantum Mechanics was developed in the early part of last century, these postulates were identified as providing the most concise formulation of the emerging theory, consistent with observation and experiment.

We shall therefore start this course by stating the postulates of quantum mechanics. You are likely to have encountered these before, perhaps written in a less formal language in terms of wave-functions or Dirac notation "ket vectors". Here, we will write these postulates in the more general language of *linear algebra*.

Linear algebra is the branch of mathematics which formalises the study of vectors, matrices and linear equations. You are likely to be familiar with many of the key concepts of linear algebra from your studies of these areas.

The advantage of the linear algebra formulation is that it applies to both wave-mechanics and Dirac formalism, allowing a unified description of both. It also allows us to more cleanly (and therefore more easily) prove general properties of the quantum systems which we study.

Here we shall begin by stating the postulates of quantum mechanics in their linear algebra formulation, and then describe, in turn, the linear algebra terms and concepts (which I have highlighted below) contained in each. There are some subtle difficulties in the formal description of continuous quantum systems (such as position or momentum of a quantum particle), so we shall focus here on discrete systems (such as atomic energy levels or spin). We shall discuss aspects of the formal treatment of continuous quantum systems in section 4, before we turn to topics which require a wave-function description, such as tunnelling and scattering.

Postulate 1: States and state Space

Every isolated physical system has associated with it a **complex inner-product space** or **Hilbert Space**. This is known as its **state space**. A complete description of the system is given by its **state vector**, $|\psi\rangle$, a **unit-normalised vector** in the state space.

Postulate 2: Evolution

The evolution of an isolated system is described the (time-dependent) Schrödinger equation,

$$i\hbar\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle \tag{1.1}$$

where H is the Hamiltonian operator, a **Hermitian operator** which specifies the full dynamics of the system.

Postulate 3: Measurement

Associated with every observable quantity is a **Hermitian operator** M with a **spectral decomposition** $M = \sum_j \lambda_j P_j$. The **eigenvalues** λ_j label the possible outcomes of the measurement. When the measurement is performed on a system in state $|\psi\rangle$, the probability that eigenvalue λ_j is returned is $\langle\psi|P_j|\psi\rangle$.

After the measurement the state of the system becomes

$$|\psi'\rangle = \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}. \tag{1.2}$$

1.3 Vector spaces

The postulates of quantum mechanics are expressed in the language of linear algebra. We shall therefore first revise the fundamentals of this branch of mathematics as we study each postulate in turn.

To understand Postulate 1, we need to introduce the abstract mathematical notion of a **vector space**. Vector spaces are an important component of linear algebra, which are widely used in theoretical physics and which many of you will

have encountered before. Here we will review the properties of vector spaces, to introduce them to those unfamiliar with them, and serve as revision for those who have.

Even if you are not familiar with this formal approach, many elements of the vector-space formalism will appear familiar to you, and you will recognize the properties and structure which you have encountered in previous quantum mechanics courses in the context of wave-functions and Dirac state-vectors.

The vector space is a formal generalisation of the familiar notion of vectors in real space. It captures many of their familiar properties of a "vector", but includes a much broader family of mathematical objects. In addition to quantum mechanics, vector spaces are also important in many theoretical areas in science, engineering and mathematics including the study of linear differential equations, Fourier analysis and coding theory.

The formal techniques in this first section may be unfamiliar to many of you. We shall be employing the techniques of formal mathematics, where objects and concepts are defined by writing down axioms, which form the basic facts about these objects which we assume to be true. The strength of the approach is that the axioms can form the basis for rigorous proofs of general properties, which will then hold for any mathematical object that satisfies the axioms.

The more abstract formulation may look fearsome at first sight, particularly if you have not encountered this kind of treatment before, but I urge you to have patience with it. In each case, examples from quantum mechanics will be provided to motivate and give intuition.

Vector spaces: An example

Before we introduce the formal definition of a vector space, let us motivate its properties with a simple example. Consider two vectors in real space $\vec{a}$ and $\vec{b}$ from the origin in a 2-D plane.

We represent such vectors as column vectors with real entries. E.g.

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad \vec{c} = \vec{a} + \vec{b} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \end{bmatrix} \tag{1.3}$$

What are the properties which characterise their "vector-character"? There are many things which one could consider; every vector has a length; between every pair of vectors, there is well defined angle; vectors can be added together and every vector can be decomposed and written as a sum of other vectors; every vector can be multiplied with a real number to produce a rescaled parallel vector.

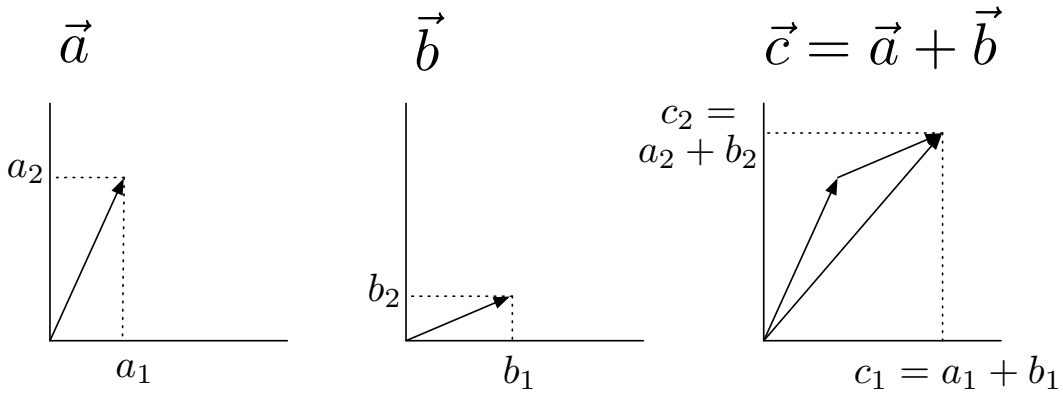

Figure 1.1: An example of 2-dimensional vectors in real space.

Perhaps surprisingly, it is the latter two observations which form the basis of the formal vector space, *vector addition* and *scalar multiplication*.

Vector addition

Vector addition is the act of *adding* pairs of vectors together to create a third vector. This property is familiar in real-space vectors and is illustrated in the figure above.

Vector addition: Axioms

We now present the formal definition of vector addition, a set of axioms which define it. Any binary operation which satisfies these axioms is an example of vector addition.

For every pair of vectors $\vec{a}$ and $\vec{b}$ there exists a vector $\vec{a} + \vec{b}$, such that the **vector addition** operation, $+$, satisfies the following characteristics:

- It is *associative*: $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$.
- It is *commutative*: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.
- There exists a zero vector: $\vec{u} + \vec{0} = \vec{u}$ for all vectors $\vec{u}$.
- Every vector has an additive inverse; for every vector $\vec{u}$ there exists a vector $-\vec{u}$ such that $\vec{u} + -\vec{u} = 0$.

These four axioms capture some of the most fundamental properties of addition. The first two reflect its behaviour under ordering and grouping; it does not matter in which order (commutativity) one adds two quantities, nor the order with which you group pairs of additions (associativity). You can verify that addition of column vectors in the above example satisfies these conditions.

Scalar multiplication

The second property that defines a vector space is scalar multiplication. This is the abstraction of the observation that every real-space vector can be multiplied with a real number to produce a rescaled copy.

Scalar multiplication: Definition

For every vector $\vec{v}$ in the vector space, there exists a vector $\lambda\vec{v}$, where λ is a real or a complex number. Scalar multiplication fulfills in addition the following properties:

- Associativity: $\lambda_1(\lambda_2\vec{a}) = (\lambda_1\lambda_2)\vec{a}$

and, when combined with scalar multiplication

- Distributivity I: $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$
- Distributivity II: $(\lambda_1 + \lambda_2)(\vec{a}) = \lambda_1\vec{a} + \lambda_2\vec{a}$

Vector Space: Definition

> Any set of objects on which vector addition and scalar multiplication, satisfying the above axioms, are defined is a **vector space**.

Vector Spaces: Examples

Column vectors

You can easily verify that column vectors with real entries (and hence vectors in real-space) satisfy the above axioms and form a formal vector space. Furthermore, if we allow column vectors with *complex entries*, this still forms a formal vector space, where the usual rules of addition and multiplication of complex numbers are employed.

One can easily verify that vector addition

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix} + \begin{bmatrix} z'_1 \\ z'_2 \\ \vdots \\ z'_{n-1} \\ z'_n \end{bmatrix} = \begin{bmatrix} z_1 + z'_1 \\ z_2 + z'_2 \\ \vdots \\ z_{n-1} + z'_{n-1} \\ z_n + z'_n \end{bmatrix} \tag{1.4}$$

satisfies the axioms of vector addition, and multiplication with a complex number

$$\lambda \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix} = \begin{bmatrix} \lambda z_1 \\ \lambda z_2 \\ \vdots \\ \lambda z_{n-1} \\ \lambda z_n \end{bmatrix} \tag{1.5}$$

satisfies the axioms of scalar multiplication.

It is worth introducing some notation here. Mathematicians usually label the the vector space of column vectors with n real entries $\mathcal{R}^n$. The vector space of column vectors with n complex entries is denoted $\mathcal{C}^n$.

Polynomial functions

The power of the vector space formalism is that it unifies a very diverse collection of different mathematical objects which share the same underlying vector structure. There are sets of objects which appear, at first sight, very different to real-space vectors, and which still form a vector space. For example, the polynomial functions form a vector space, with vector addition and scalar multiplication defined as one would expect.

To see this, consider a pair of polynomials, e.g. $p_1(x) = x^2$ and $p_2(x) = x^3$. We can combine them using addition of functions and scalar multiplication to form another polynomial $p_3(x) = ap_1(x) + bp_2(x) = ax^2 + bx^3$. The null vector in this vector space is the constant function $f(x) = 0$. As an exercise, verify that the axioms of a vector space are satisfied.

Wave-functions and Dirac state vectors

Since functions and column vectors form a vector space, it should not be a surprise that wave-functions and Dirac ket-vectors also form a vector space, as Postulate 1

demands. In a sense the vector space requirement for quantum states embodies the "principle of superposition" at a more formal level. Superposition is represented by the vector addition and scalar multiplication of the vector space.

We have now defined formally the vector space, but there is additional formalism in Postulate 1, which we have not introduced yet. Postulate 1 includes terms such as "inner product space" and "unit-normalised vector". To understand these concepts we need to introduce the *vector norm* and the *inner product*, which we shall do in the next section.

1.4 A notion of length: the vector *norm*

The definition of vector spaces introduced above, only refers to two properties of real space vectors (vector addition and scalar multiplication) out of the many we identified. So far these abstract spaces do not possess analogues of vector properties such as length or angles. We need to introduce further structure to define these.

The first properties we will consider is length. We would like to introduce an analogue of length for the abstract vector spaces introduced above. It is not immediately obvious how one would do this. For example, what would be the "length" of a complex vector? The "length" of a polynomial function?

As before, we need to capture the essential properties of length by a set of axioms. The quantity which we will define this way, is called the *vector norm*. As we shall see, it is in fact, a much more general concept than length.

Vector norm: Definition

Let us first write down the axioms which a norm must satisfy before giving some examples to motivate them.

> A vector norm is a function $||\vec{a}||$ which associates with every vector $\vec{a}$ a non-negative real number. The norm satisfies the following properties:
>
> 1. $||\lambda\vec{a}|| = |\lambda|\ ||a||$. "Scaling" behaviour
> 2. $||\vec{b}|| = 0$ if and only if $\vec{b}$ is the zero vector.
> 3. $||\vec{a} + \vec{b}|| \leq ||\vec{a}|| + ||\vec{b}||$, the *triangle inequality*.

The first two axioms are rather natural requirements for an analogue of "length", but the third deserves some discussion. To see why an analogue of length should satisfy this consider the following figure

$$\vec{c} = \vec{a} + \vec{b}$$

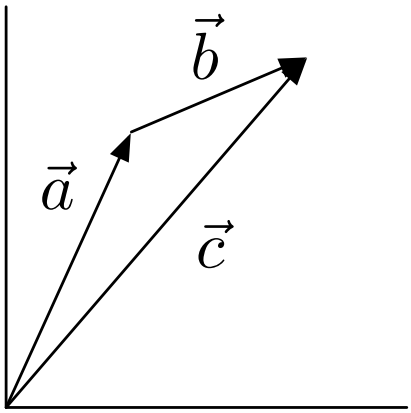

Figure 1.2: Illustration of the "triangle inequality" for the lengths of real vectors.

You can see that the length of $\vec{c}$ can never be greater than the combined lengths of $\vec{a}$ and $\vec{b}$ though will often be less than this maximum. The triangle inequality demands that all vector norms share this characteristic.

Any vector space where a norm is defined is called a *normed vector space*.

Vector norm: Examples

Euclidean norm

The length of a vector in real space $\vec{a}$ is (as one would expect) a vector norm.

$$||\vec{a}|| = \sqrt{\vec{a} \cdot \vec{a}} = \sqrt{\sum_i a_i^2} \tag{1.6}$$

It is straightforward to confirm that this length does satisfy the first two norm axioms, and by considering Figure 1.2 or otherwise one can confirm that the trian-

gle inequality is always satisfied. This norm is often referred to as the *Euclidean norm*.

Euclidean norm for complex vectors

Vectors with complex entries do not have a "length" in the usual sense, and simply utilising equation (1.6) for complex valued vectors will not be sufficient, since non-real entries can lead to the sum in equation (1.6) to be complex, while a norm must be a real quantity.

This can be resolved by adopting the following definition for the Euclidean norm for a complex vector $\vec{z}$:

$$||\vec{z}|| = \sqrt{\sum_i |z_i|^2} \tag{1.7}$$

Via this simple modification of equation (1.6) (which includes (1.6) as a special case when all elements are real) we see that the norm axioms are satisfied.

Norms for functions

How can one define a norm for a function? If the norm is to be interpreted as an analogue of the length of a vector, it should, in some sense reflect the "size" of the function. The null vector in the vector space for polynomial functions was the $f(x) = 0$ function, and thus any norm for this function must be zero.

One way to gain an intuition for this is to consider the function as a continuous version of a vector, where the discrete vector elements have been replaced by infinitessimal points:

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix} \rightarrow \begin{bmatrix} \vdots \\ z(x) \\ \vdots \end{bmatrix} \tag{1.8}$$

Under that analogy, it is natural to replace the sums in the above discrete norms, with integrals, and define a norm on the space of functions analogously. In fact, the integral can, indeed, be used to define norms on the vector space of

polynomials, such as the following natural generalisation of the Euclidean norm.

$$||f|| = \sqrt{\int |f(x)|^2 dx} = \sqrt{\int f(x)^* f(x) dx} \tag{1.9}$$

There are several important points to note. Defined in this way, the norm depends upon the range of integration, thus this must be clearly stated. Secondly, for certain functions, the integral may diverge. Such functions do not possess a finite norm. This will be an important observation to be taken into account when constructing wave-functions for physical states. This, and other issues regarding the formal treatment of wave-functions, will be discussed in more depth in part 4.

Norms in quantum mechanics

In quantum mechanics, the norm plays an important role, since it is closely related to probability. However, before we introduce norms for quantum states, we will first introduce a closely related concept, the *inner product*.

1.5 Abstracting "angle": the inner-product

An important property of pairs of real space vectors is the angle between them. The angle between real vectors $\vec{a}$ and $\vec{b}$ is related to their inner product (also called "dot product" or "scalar product") by the following identity.

$$\vec{a} \cdot \vec{b} = \cos(\theta) ||\vec{a}|| \; ||\vec{b}|| \tag{1.10}$$

where $||\vec{a}||$ is the Euclidean norm.

We see that angle, norm and inner product are closely related for real spaces. In the abstract setting, it is the inner product which is most convenient to generalise.

Similar to the norm and the vector space, inner products for abstract vector spaces are defined in terms of a set of axioms which they must satisfy. Again, the axioms are chosen to capture the most important properties of the inner product in real space.

Inner product: Definition

An inner product between two vectors is a function, often written $(\vec{a}, \vec{b})$ which associates with every pair of vectors $\vec{a}$ and $\vec{b}$ a complex number.

The inner product satisfies the following properties:

1. $(\vec{a}, \lambda\vec{b}) = \lambda(\vec{a}, \vec{b})$: Linearity in the second argument
2. $(\vec{a}, \vec{a}) \geq 0$: Positive-definiteness. The inner product of a vector with itself is non-negative, and equal to zero if and only if $\vec{a}$ is the null vector.
3. $(\vec{a}, \vec{b}) = (\vec{b}, \vec{a})^*$ - Conjugate symmetry: interchanging the order of the vectors in the inner product, leads to complex conjugation of the inner product

One can verify immediately that these axioms are satisfied by the real-space "dot product". This choice of axioms may seem at first sight strange, especially axioms 2 and 3, but they are chosen for good reason. As we shall see below, they are defined as they are to preserve an important relationship between the inner product and a norm.

The linearity in Axiom 1 is specific to the second argument. In fact, the inner product is not linear in the first argument, since, by combining axioms 1 and 3 we see:

$$\begin{aligned} (\lambda\vec{a}, \vec{b}) &= (\vec{b}, \lambda\vec{a})^* && \text{By axiom 3} \\ &= [\lambda(\vec{b}, \vec{a})]^* && \text{By axiom 1} \\ &= \lambda^*(\vec{b}, \vec{a})^* \\ &= \lambda^*(\vec{a}, \vec{b}) && \text{By axiom 3} \end{aligned} \tag{1.11}$$

We call this "anti-linearity" in the first argument. If the first vector is multiplied by a scalar λ, the inner-product is multiplied by its complex conjugate λ^*.

A vector space with an inner product defined is called an *inner product space*.

Inner products in quantum mechanics

In quantum mechanics, the inner product between two states is sometimes called the "overlap". In the Dirac formalism it is written

$$(|\psi\rangle, |\phi\rangle) = \langle\psi|\phi\rangle \tag{1.12}$$

For wave-functions it is defined

$$(\psi(x), \phi(x)) = \int \psi(x)^* \phi(x) dx \tag{1.13}$$

You should verify as an exercise that these inner products do satisfy the above axioms.

1.6 Norm from inner product

The definition of the inner product is perhaps the least intuitive of the definitions we have seen so far. While one can readily verify that a real-space dot-product satisfies the above axioms, it is, at first glance, hard to see why these particular axioms were chosen. In fact, the axioms take this form to ensure that one important property always holds, namely that the inner product itself defines a norm.

> **Theorem** Given a vector $\vec{a}$ in an inner product space, the quantity $\sqrt{(\vec{a}, \vec{a})}$ is a norm of $\vec{a}$.

The proof of this theorem can be found in linear algebra textbooks. Instead of a detailed proof, we shall illustrate the theorem via some examples.

Norm from inner product: Example

As an example of this, let's verify that the standard inner product for real-space vectors (the dot product) does indeed define a norm, and that that norm is the Euclidean length.

$$\sqrt{\vec{a} \cdot \vec{a}} = \sqrt{\sum_j a_j a_j} = \sqrt{\sum_j a_j^2} = ||a|| \tag{1.14}$$

Norm from inner product for quantum states

The norm which we employ for quantum states is the one derived from the inner products in equations (1.12) and (1.13),

$$|| \, |\psi\rangle || = \sqrt{\langle\psi|\psi\rangle} \tag{1.15}$$

$$||\phi(x)|| = \sqrt{\int \phi(x)^* \phi(x) dx} \tag{1.16}$$

Normalisation

In quantum mechanics, we say a state vector is *normalised* when it has unit norm, i.e. $\sqrt{\langle\psi|\psi\rangle} = 1$. Note that this implies that $\langle\psi|\psi\rangle = 1$. To simplify the calculation of probabilities in quantum mechanics we usually work with states which are normalised.

1.7 Basis, linear independence and dimension

We have now introduced most of the linear algebraic concepts which appear in Postulate 1. To conclude our introduction to vector spaces, we shall survey some important properties of vector spaces which are important in quantum mechanics.

One of the most important properties of a vector space is its **dimension**. For real-space vectors, the idea of dimension is intuitive, i.e. the space of n- entry complex column vectors $\mathcal{C}^n$ is n-dimensional. We would like to define dimension for abstract vector spaces. To do so, we need to introduce **spanning sets**, **linear dependence** and **basis sets**.

Spanning set

We know that in real space, every vector can be expressed as a weighted sum of a set of other vectors. For example, in Figure 1.1, $\vec{c}$ is expressed as the sum $\vec{a} + \vec{b}$, but similarly we could write, $\vec{a} = \vec{c} - \vec{b}$, etc.

In particular, we also often express real-space vectors as a weighted sum of a standard set of basis vectors. E.g. in 3-dimensional space ($\mathcal{R}^3$):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x\vec{i} + y\vec{j} + z\vec{k} = x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{1.17}$$

Any vector in $\mathcal{R}^3$ can be expressed as a sum of the three basis vectors $\vec{i}$, $\vec{j}$, $\vec{k}$. We say that these basis vectors **span** the space $\mathcal{R}^3$. In general, a **spanning set** is

any set of vectors $\vec{w_j}$, as a weighted sum of which, any vector $\vec{v}$ in the space can be represented i.e.

$$\vec{v} = \sum_j \alpha_j \vec{w}_j \tag{1.18}$$

Example: Spanning set

A vector space will have many different spanning sets. For example, in problem sheet 1 you will show that the following set is also a spanning set for $\mathcal{R}^3$ and $\mathcal{C}^3$:

$$\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \qquad \begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \tag{1.19}$$

Note that the spaces $\mathcal{R}^n$ and $\mathcal{C}^n$ can be built from the same spanning set, which we call the standard basis,

$$\vec{e_1} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \vec{e_2} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \qquad \vec{e_n} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{1.20}$$

The two spaces differ since $\mathcal{C}^n$ allows scalar multiplication over complex numbers whereas for $\mathcal{R}^n$ scalar multiplication is restricted to the reals. We say that $\mathcal{C}^n$ is a vector space *over the complex numbers*, or more compactly, that $\mathcal{C}^n$ is a *complex vector space*.

Linear Dependence and Linear Independence

If you take a spanning set, and add an additional vector to the set, the new larger set will also span the space. That means that there must exist spanning sets which have more than the minimal necessary vectors. It would be desirable to use the smallest spanning set, but how can we verify that the spanning set we have is the most compact? This can be achieved by considering the *linear independence* of the vectors in the set.

Linear Dependence and Linear Independence: Definition

We call a set of non-zero vectors $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n$ **linearly dependent** when there exists a set of (non-zero) scalars a_j such that

$$a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_n\vec{v_n} = 0. \tag{1.21}$$

A set of vectors is **linearly independent** if it is not linearly dependent.

Linear Dependence and Linear Independence: Example

An example of a linear independent set is the set of standard basis vectors for $\mathcal{C}^3$:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{1.22}$$

You should verify as an exercise that these are indeed linearly independent.

Basis vectors

Given any spanning set with linearly dependent vectors, it is always possible to reduce the set by removing vectors until the set left over is linearly independent. We shall not prove this here, but illustrate it by example.

The following three vectors in $\mathcal{R}^2$ form a spanning set:

$$\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \vec{w}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \vec{w}_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \tag{1.23}$$

and hence any vector in $\mathcal{R}^2$ can be written $\vec{v} = a_1\vec{w}_1 + a_2\vec{w}_2 + a_3\vec{w}_3$. However, this set is not linearly independent since:

$$2\vec{w}_1 - \vec{w}_2 + \vec{w}_3 = 0. \tag{1.24}$$

We can rewrite equation (1.24) as

$$\vec{w}_2 = 2\vec{w}_1 + \vec{w}_3 \tag{1.25}$$

and then substitute equation (1.25) into $\vec{v} = a_1\vec{w}_1 + a_2\vec{w}_2 + a_3\vec{w}_3$ to write $\vec{v}$ in terms of $\vec{w}_1$ and $\vec{w}_3$ only:

$$\vec{v} = (a_1 + 2a_2)\vec{w}_1 + (a_2 + a_3)\vec{w}_3 \tag{1.26}$$

We see that the $\vec{w}_1$ and $\vec{w}_3$ alone form a spanning set. You can verify that this set is linearly independent.

Basis: definition

Linear independent spanning sets have the minimum number of elements, and play a very important role in linear algebra.

> We call a linearly independent spanning set a **basis**.

The identification of a basis set for a space, also identifies that space's dimension.

Dimension: definition

One can prove that for any vector space *all* basis sets for that space share the same number of elements.

> The number of elements of a basis set for a vector space is called its **dimension**.

The dimension d corresponds to the number of free parameters for a generic vector in the space, i.e.

$$\vec{v} = \sum_{j=1}^{d} \alpha_j \vec{w}_j \tag{1.27}$$

for basis vectors $\vec{w}_j$ for j from 1 to d.

Dimension: example

The dimension of $\mathcal{R}^3$ is three, since the set of vectors (1.22) form a basis. In general, the dimension of $\mathcal{R}^n$ and $\mathcal{C}^n$ is n, as one can readily verify by considering the standard basis vectors for each.

Dimension: example 2

In quantum mechanics, the state-space of a spin $n/2$ particle is $n+1$. In particular, a spin $1/2$ particle is 2-dimensional. A basis set is given by spin up $|\uparrow\rangle$ and spin down $|\downarrow\rangle$ states.

Dimension: example 3

The dimension of a vector space does not have to be finite. For example, the vector space of polynomials has infinite dimension, since the functions $f_j(x) = x^k$ form a basis for this space. You should verify this for yourself, as an exercise.

1.8 Orthogonality

The vectors in the standard basis set in real space are perpendicular or **orthogonal**. We would like to generalise orthogonality to general vector spaces. This can be achieved by considering the following observation. In real space, two vectors are orthogonal if and only if their inner product is zero. We generalise this directly.

Orthogonality: Definition

> In an inner product space, two vectors are **orthogonal** if and only if their **inner product** is zero.

Orthogonality: Examples

You have encountered this kind of orthogonality in quantum mechanics. We call two Dirac vectors orthogonal when

$$\langle\psi|\phi\rangle = 0 \tag{1.28}$$

and two wave-functions are orthogonal when

$$\int \psi(x)^* \phi(x) dx = 0 \tag{1.29}$$

Orthonormal basis and the closure relation

A basis which is orthogonal and normalised is called an **orthonormal basis**. Orthonormal bases have so many useful properties that we use them almost invariably in quantum mechanics. The orthonormality condition $\langle\phi_j|\phi_k\rangle = \delta_{j,k}$ is employed frequently to simplify calculations.

Another important property of orthonormal bases is the *closure relation*, which will be introduced below in section 1.15.

1.9 Linear functionals and the dual space

In Dirac notation, the inner product between two state vectors $|\psi\rangle$ $|\phi\rangle$ is written $(|\psi\rangle, |\phi\rangle) = \langle\psi|\phi\rangle$. This notation is suggestive of a multiplication between two objects $\langle\psi|$ and $|\phi\rangle$. It is likely that you have encountered so-called "bra-vectors" such as $\langle\psi|$, when you first learnt about Dirac notation. In this section, we shall learn the mathematical description of these objects.

In fact, they have two equivalent definitions, as (rather fearsome sounding) **linear functionals** and as elements of a **dual** vector space.

Linear functional: Definition

The linear functional has a simple definition.

> A *functional* $f(\vec{v})$ is a function which acts on a vector and which returns a scalar. A *linear* functional is a functional which fulfils the following requirement:

$$f(\alpha_1\vec{v_1} + \alpha_2\vec{v_2}) = \alpha_1 f(\vec{v_1}) + \alpha_2 f(\vec{v_2}) \tag{1.30}$$

Linear functional: Example

The bra $\langle\psi|$ represents a linear functional $f_\psi(|\phi\rangle)$ since it acts on vector $|\phi\rangle$, returns a scalar $\langle\psi|\phi\rangle$ and satisfies

$$f_\psi(\alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle) = \langle\psi|\left(\alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle\right) = \alpha_1\langle\psi|\phi_1\rangle + \alpha_2\langle\psi|\phi_2\rangle \tag{1.31}$$

due to the axioms of the inner product. Since we used the axioms of the inner product to prove this property, this holds true for *any* inner product space. All inner products define a linear functional.

Dual space

The above argument shows us that we can associate a "bra" linear functional $\langle\psi|$ with every vector $|\psi\rangle$ in the vector space. It is thus natural to ask whether the bra objects can themselves be thought of vectors in their own right?

Dual space: definition

> For any vector space V, the set of linear functionals acting on that space is a vector space V^* called the **dual** of V.

It is left as an exercise to verify that this is indeed a vector space.

Examples: Column vectors

The dual of the set of complex column vectors is the set of row vectors. To see this, consider an element of the space of two-element column vectors, $\mathcal{R}^2$;

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{1.32}$$

Its dual is the row vector

$$\vec{v}^* = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \tag{1.33}$$

To see that this is true, remember that we can write the inner product between $\vec{v}$ and a second vector $\vec{w}$ as

$$\vec{v} \cdot \vec{w} = \sum_j v_j w_j = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \tag{1.34}$$

where the multiplication between the column and row vectors is now standard matrix multiplication.

In n-dimensional quantum mechanics, a state $|\psi\rangle$ can be represented by an n-element complex column vector.

$$|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{1.35}$$

The corresponding dual vector, bra $\langle\psi|$ is the conjugate transpose vector,

$$\langle\psi| = \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} \tag{1.36}$$

You can verify that this dual vector does indeed have all the required properties of a linear functional.

1.10 Hilbert Space

The state space of quantum mechanics is a mathematical object known as **a Hilbert space**. In finite dimensions, the definition is simple.

Hilbert Space: definition for finite dimensions

An inner product space of finite dimension is a *Hilbert space*.

In an infinite dimensional system, the definition is much more technical, due to some issues which arise with convergence in the infinite-dimensional case. This infinite-dimensional definition (which is what a *mathematician* would mean by the term Hilbert space) will not be covered in this course. You will find a clear discussion of infinite-dimensional Hilbert spaces (and their definition) in chapter 4 of "Quantum Theory" by Asher Peres. It has become common for physicists to use the term Hilbert space for finite dimensional state spaces (when in fact the term "inner product space" would suffice!) and you will find it commonly used in advanced textbooks and research literature.

1.11 Postulate 1 and Dirac notation

We have now covered all of the mathematical background needed to understand Postulate 1.

Postulate 1: States and state Space

Every isolated physical system has associated with it a **complex inner-product space** or **Hilbert Space**. This is known as its **state space**. A complete description of the system is given by its **state vector**, $|\psi\rangle$, a **unit-normalised vector** in the state space.

This compact description allows all of the mathematical machinery, norms, inner products, dual spaces, introduced above to be incorporated into the definition of a

quantum state vector. The Dirac notation will be familiar to you, but it is worth revising briefly in terms of its correspondence with the linear algebraic concepts introduced above.

Dirac notation in linear algebraic terms

Notation	Description
$\|\psi\rangle$	Vector in a state space (also known as a *ket*)
$\langle\psi\|$	Dual vector to $\|\psi\rangle$ (also known as a *bra*)
$\langle\phi\|\psi\rangle$	Inner product between vectors $\|\phi\rangle$ and $\|\psi\rangle$

From this point onwards we shall write all further discussion of linear algebraic concepts using Dirac notation.

Postulate 1: Example: Spin half particle

The spin-half particle will play an important role in this course, for a number of reasons. Spin-half particles, including the electron, proton and neutron, play a central role in Physics, and the simplicity of their state space makes it an ideal example of the quantum formalism.

The state space for a spin-half particle is 2-dimensional, with basis states $|1/2, +1/2\rangle = |\uparrow\rangle$ and $|1/2, -1/2\rangle = |\downarrow\rangle$ spanning a 2-dimensional state space. An arbitrary state can be written:

$$|\psi\rangle = \alpha_\uparrow|\uparrow\rangle + \alpha_\downarrow|\downarrow\rangle \tag{1.37}$$

where, to preserve unit-normalisation $|\alpha_\uparrow^2| + |\alpha_\downarrow^2| = 1$.

We shall now turn to Postulates 2 and 3, which require us to define and represent *operators* which act on this state space.

1.12 Subspaces

We know, from our own experience with real space, that vector spaces can contain smaller spaces inside them. For example, a two-dimensional plane has all the properties of a (two-dimensional) vector space. We call the plane a sub-space of three-dimensional space.

Subspace: Definition

Given a vector space V, a sub-space W is a space containing some of the vectors of V and satisfying all of the properties of a vector space in its own right. In particular, the sub-space must contain the zero vector, and vectors in a sub-space must be closed under addition and scalar multiplication, i.e. given any two vectors in W, v_1 and v_2, we know that $\alpha v_1 + \beta v_2$ is also in W. To denote that W is a subspace of V we write $W \subset V$.

For an example, the vectors in a 2-dimensional plane form a 2-dimensional subspace in three dimensional space. E.g. vectors

$$\vec{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \vec{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{1.38}$$

form an orthonormal basis for the subspace consisting of the $z = 0$ plane in three-dimensional space.

Sub-spaces are important in quantum mechanics in a number of contexts, for example, in quantum measurement.

1.13 Column vector representation of quantum states

We have introduced an abstract vector space representation for quantum states, but often for calculations and to help our intuition it is useful to have a representation which is more concrete. Here we shall see that any state in a (finite) n-dimensional space may be represented by a n-element complex column vector (i.e. in the vector space $\mathcal{C}^n$.)

Imagine that we have identified an orthonormal basis $|\phi_j\rangle$ for $j = 1$ to n for

the vector space of our quantum state. We can thus write any state, $|\psi\rangle$

$$|\psi\rangle = \sum_{j=1}^{n} \alpha_j |\phi_j\rangle \tag{1.39}$$

Consider now a column-vector of coefficients α_j

$$\vec{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \tag{1.40}$$

We can use this column vector to represent $|\psi\rangle$ and calculate any of the properties of states in this space.

For example, the norm of $|\psi\rangle$

$$\begin{aligned} ||\,|\psi\rangle|| = \sqrt{\langle\psi|\psi\rangle} &= \sqrt{\sum_{j,k} \left(\alpha_j^* \langle\phi_j|\right) \left(\alpha_k |\phi_k\rangle\right)} \\ &= \sqrt{\sum_{j,k} \alpha_j^* \alpha_k \langle\phi_j|\phi_k\rangle} = \sqrt{\sum_{j,k} \alpha_j^* \alpha_k \, \delta_{j,k}} \\ \sqrt{\sum_{j,k} \alpha_j^* \alpha_j} &= ||\vec{\alpha}|| \end{aligned} \tag{1.41}$$

is equal to the complex Euclidean norm of the column vector, where we used the fact that since $|\phi_j\rangle$ is an orthonormal basis $\langle\phi_j|\phi_k\rangle = \delta_{j,k}$.

And more generally, the inner product between two states $|\psi\rangle = \sum_{j=1}^{n} \alpha_j |\phi_j\rangle$ and $|\chi\rangle = \sum_{j=1}^{n} \beta_j |\phi_j\rangle$,

$$\begin{aligned} \langle\psi|\chi\rangle &= \sum_{j,k} \left(\alpha_j^* \langle\phi_j|\right) \left(\beta_k |\phi_k\rangle\right) \\ &= \sum_{j,k} \alpha_j^* \beta_k \langle\phi_j|\phi_k\rangle = \sum_{j,k} \alpha_j^* \beta_k \, \delta_{j,k} \\ &= \sum_{j,k} \alpha_j^* \beta_j = (\vec{\alpha}, \vec{\beta}), \end{aligned} \tag{1.42}$$

is equal to the inner product between column vectors $\vec{\alpha}$ and $\vec{\beta}$.

The bra-vector $\langle\psi|$ is represented by the dual of $\vec{\alpha}$, which (as shown above) is the transpose complex conjugate of $\vec{\alpha}$ the row vector

$$\begin{bmatrix}\alpha_1^* & \alpha_2^* & \cdots & \alpha_n^*\end{bmatrix}. \tag{1.43}$$

Note that in this vector representation the orthonormal basis states $|\phi_j\rangle$ are represented by the standard basis for $\mathcal{C}^n$

$$\vec{e_1} = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix} \qquad \vec{e_2} = \begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix} \qquad \vec{e_n} = \begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}. \tag{1.44}$$

E.g, for a spin half particle, we can make the identification:

$$|\uparrow\rangle \cong \begin{bmatrix}1\\0\end{bmatrix} \qquad |\downarrow\rangle \cong \begin{bmatrix}0\\1\end{bmatrix} \tag{1.45}$$

and can write an arbitrary state

$$|\psi\rangle = \alpha_1|\uparrow\rangle + \alpha_2|\downarrow\rangle \cong \begin{bmatrix}\alpha_1\\\alpha_2\end{bmatrix} \tag{1.46}$$

Note that I have used the symbol $\cong$ instead of equals here. This is rather pedantic notation but emphasises that the column vector is a representation of $|\uparrow\rangle$ it is not "equal" to (i.e. the same mathematical object as) $|\uparrow\rangle$. Most textbooks will not be so careful in their notation and will use the $=$ sign in this context.

Below we shall also see that just as states can be represented by column vectors, abstract operators may be represented by matrices.

1.14 Operators

Postulates 2 and 3 refer to the time-evolution and measurement of quantum states. In both cases *linear operators* are the mathematical objects which capture the transformations undergone by states as they evolve and are measured.

Linear operator: Definition

An operator $\hat{O}$ acting on a vector space is a function which acts on a vector (e.g. $|\psi\rangle$) in the space and returns a vector in the same space (e.g. $|\psi'\rangle$)

$$|\psi'\rangle = \hat{O}|\psi\rangle \tag{1.47}$$

We often write operators with a "hat" $\hat{O}$ to differentiate them from classical variables and scalar quantities. I will not always do so, but will try to use the "hat" wherever an operator could be confused with a classical value. Note that conventionally we do not write an operator acting on a vector as a function, $\hat{O}(|\psi\rangle)$, but as a multiplication $\hat{O}|\psi\rangle$.

> A linear operator $\hat{A}$ is an operator which acts linearly, meaning that the following condition holds
>
> $$\hat{A}[\alpha_1|\psi_1\rangle + \alpha_2|\psi_2\rangle] = \hat{A}\alpha_1|\psi_1\rangle + \hat{A}\alpha_2|\psi_2\rangle \tag{1.48}$$

Linear operators play a very important role in quantum mechanics, particularly in time evolution and measurement.

They have many important properties which we shall describe in the following sections. Note that equation (1.48) can be used to show that the sum of two linear operators is itself a linear operator.

Operator multiplication and the Commutator

Operators may be applied sequentially. Consider the sequence of operators $\hat{O}_1$, $\hat{O}_2$, $\hat{O}_3$, etc. If we apply these in turn to a state $|\psi\rangle$ we obtain

$$\hat{O}_3\hat{O}_2\hat{O}_1|\psi\rangle. \tag{1.49}$$

This notation means first apply $\hat{O}_1$ to state $|\psi\rangle$ then apply O_2 then O_3. Thus to read the "time-order" of the operators we read right-to-left not left-to-right. We can also consider $\hat{O}_3\hat{O}_2\hat{O}_1$ as representing the product of the three operators. Recall that unlike scalar multiplication, the order in which we write operators in an equation or expression is very important.

We use the *commutator* to capture this. For a pair of operators $\hat{A}$ and $\hat{B}$ we write

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{1.50}$$

When the commutator of two operators is zero, we say the operators *commute*. The order of commuting operators may be interchanged, since $\hat{A}\hat{B} - \hat{B}\hat{A} = 0$ implies $\hat{A}\hat{B} = \hat{B}\hat{A}$.

1.15 Operators from states and the closure relation

We saw above that if we take the product of a state $|\psi\rangle$ and a dual $\langle\phi|$ we obtain the inner product $\langle\phi|\psi\rangle$. Consider what we obtain though if we take this product in the opposite order $|\psi\rangle\langle\phi|$. The product in this ordering is *not* an inner product, but something else. In fact, $|\psi\rangle\langle\phi|$ (sometimes called an outer product) is a linear operator which transforms a state $|\chi\rangle$ into state $|\chi'\rangle$ as follows,

$$|\chi'\rangle = (|\psi\rangle\langle\phi|)\,|\chi\rangle = |\psi\rangle\langle\phi|\chi\rangle = \langle\phi|\chi\rangle|\psi\rangle. \tag{1.51}$$

We see that $|\psi\rangle\langle\phi|$ has converted state $|\chi\rangle$ into state $|\psi\rangle$ multiplied by the scalar prefactor $\langle\phi|\chi\rangle$, and is thus an operator. You can confirm that this is in fact a linear operator.

Closure relation

Let us now consider an operator, constructed in this manner from orthonormal basis states $|\phi_j\rangle$ and their dual $\langle\phi_j|$

$$\sum_j |\phi_j\rangle\langle\phi_j|. \tag{1.52}$$

Let us consider how this operator acts on an arbitrary state $|\psi\rangle = \sum_k \alpha_k|\phi_k\rangle$,

$$\begin{aligned}\left(\sum_j |\phi_j\rangle\langle\phi_j|\right)|\psi\rangle &= \sum_j |\phi_j\rangle\langle\phi_j| \sum_k \alpha_k|\phi_k\rangle \\ &= \sum_{j,k} \alpha_k\langle\phi_j|\phi_k\rangle|\phi_j\rangle = \sum_{j,k} \alpha_k\delta_{j,k}|\phi_j\rangle \\ &= \sum_k \alpha_k|\phi_k\rangle = |\psi\rangle.\end{aligned} \tag{1.53}$$

We see that this operator leaves the arbitrary state $|\psi\rangle$ unchanged. This means that it leaves all states unchanged, in other words, this operator is the *identity operator*,

$$\sum_j |\phi_j\rangle\langle\phi_j| = \mathbb{1}. \tag{1.54}$$

This expression is called the **closure relation** or, more commonly, the **resolution of the identity**. It is simple and a direct consequence of the orthonormality of

$|\phi_j\rangle$. This relation is extremely useful, since the identity operator can be inserted at will into any expression. We shall use this trick several times in the remainder of this course.

1.16 Matrix representations of linear operators

We have seen how states can be represented by column vectors. Similarly, linear operators can be represented by matrices. In the matrix representation, multiplication of operators, and applying an operator to a state, are both represented by the usual rules of matrix multiplication.

We find the matrix representation of operator A by exploiting the resolution of the identity for basis states $|\phi_j\rangle$

$$\begin{aligned} A = \mathbb{1}\, A\, \mathbb{1} &= \sum_{j,k} |\phi_j\rangle\langle\phi_j|\, A\, |\phi_k\rangle\langle\phi_k| \\ &= \sum_{j,k} \langle\phi_j|A|\phi_k\rangle |\phi_j\rangle\langle\phi_k| \\ &= \sum_{j,k} A_{j,k} |\phi_j\rangle\langle\phi_k| \end{aligned} \tag{1.55}$$

where we call $A_{j,k} = \langle\phi_j|A|\phi_k\rangle$ a *matrix element* of A. Note how the Dirac notation makes it easy to see that this is a scalar. In Dirac notation, any object which opens with $\langle$ and closes with $\rangle$ is a scalar.

If we consider how A acts on a general state $|\psi\rangle = \sum_m \alpha_m |\phi_m\rangle$ we see why $A_{j,k}$ define the matrix representation of A.

$$\begin{aligned} A|\psi\rangle &= \sum_{j,k} A_{j,k} |\phi_j\rangle\langle\phi_k| \sum_m \alpha_m |\phi_m\rangle \\ &= \sum_{j,k,m} A_{j,k}\alpha_m |\phi_j\rangle\langle\phi_k|\phi_m\rangle \\ &= \sum_{j,k,m} A_{j,k}\alpha_m |\phi_j\rangle\delta_{k,m} \\ &= \sum_{j,m} A_{j,m}\alpha_m |\phi_j\rangle \\ &= \sum_j \beta_j |\phi_j\rangle \end{aligned} \tag{1.56}$$

where $\beta_j = \sum_m A_{j,m}\alpha_m$.

Recall that matrix multiplication is written, at the level of matrix elements as $[AB]_{j,k} = \sum_m A_{j,m}B_{m,k}$. Hence $\vec{\beta} = A\vec{\alpha}$. You can see then that the vector representation $\vec{\alpha}$ of the state is transformed by matrix multiplication through the matrix A with elements $A_{j,k} = \langle\phi_j|A|\phi_k\rangle$.

1.17 Hermitian Conjugate

You will recall that Hermitian conjugates or adjoints play an important role in quantum mechanics. Here, we briefly revise their definition: Consider an operator sandwiched between a bra and a ket

$$\langle\phi|A|\psi\rangle. \tag{1.57}$$

We can consider this as the product of dual vector $\langle\phi|$ and the transformed ket $A|\psi\rangle$,

$$\langle\phi|A|\psi\rangle = \langle\phi|\Big(A|\psi\rangle\Big). \tag{1.58}$$

Alternatively, we may wish to consider it as the product of the transformed dual vector $\langle\phi|A$ and the ket $|\psi\rangle$,

$$\langle\phi|A|\psi\rangle = \Big(\langle\phi|A\Big)|\psi\rangle \tag{1.59}$$

where the operator A is now acting from the right on dual vector $\langle\phi|$.

So far, however, we have not made precise what an operator "acting from the right" on a dual vector means. The *Hermitian conjugate* or *adjoint* allows us to do this. The Hermitian conjugate of operator A, denoted $A^\dagger$, is usually defined in standard inner product notation as follows:

$$(A^\dagger|\psi\rangle, |\phi\rangle) = (|\psi\rangle, A|\phi\rangle) \tag{1.60}$$

This definition is equivalent to the following, more convenient definition, expressed using Dirac notation:

$$(\langle\phi|A^\dagger|\psi\rangle)^* = \langle\psi|A|\phi\rangle \tag{1.61}$$

where we used axiom 1 for the inner product $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$.

This definition means that the dual of the vector $A|\psi\rangle$ is equal to $\langle\psi|A^\dagger$. Hence, one way of considering the Hermitian conjugate is that it converts an operator acting on kets to an equivalent operator acting on bras. It is conventional to also define Hermitian conjugation for vectors and scalars, i.e. $|\psi\rangle^\dagger = \langle\psi|$, $\alpha^\dagger = \alpha^*$.

In matrix representations, Hermitian conjugation always corresponds to taking the complex conjugate and transpose.

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^\dagger = \begin{bmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{bmatrix} \qquad \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}^\dagger = \begin{bmatrix} \alpha_1^* \\ \alpha_2^* \end{bmatrix} \tag{1.62}$$

When we take the Hermitian conjugate of a product of operators (and vectors) it reverses the order of multiplication. This is a consequence of equation (1.61). Thus,

$$\langle\phi|\hat{A}\hat{B}|\psi\rangle^\dagger = \langle\psi|\hat{B}^\dagger\hat{A}^\dagger|\phi\rangle \tag{1.63}$$

There are some subtle issues regarding adjoints and Hermitian conjugates for infinite-dimensional systems (in which they have a different definition and are not always equal), but for finite-dimensional systems, and all examples of infinite-dimensional systems which you will encounter in this course, the two concepts can be considered identical.

1.18 Unitary Operator

The *inverse* $\hat{A}^{-1}$ of an operator $\hat{A}$ is the operator which, when multiplied by $\hat{A}$ gives the identity, i.e. $\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \mathbb{1}$.

Definition

> An operator U is **unitary** if its Hermitian conjugate is its inverse, e.g. if $UU^\dagger = U^\dagger U = \mathbb{1}$.

Unitary operators play a very important role in quantum mechanics and have a number of useful properties. They

- are norm-preserving, that is $||U|\psi\rangle|| = ||\,|\psi\rangle||$.
- preserve the inner product between pairs of vectors $(U|\phi\rangle, U|\psi\rangle) = \langle\phi|U^\dagger U|\psi\rangle = \langle\phi|\psi\rangle$.

- have eigenvalues of the form $e^{i\theta}$.

The first two properties mean that unitary operators can be employed to *change basis*, i.e. given two orthonormal basis sets for the same space $\{|\phi_i\rangle\}$ and $\{|\psi_i\rangle\}$ one can always find a unitary operator U such that $U|\phi_i\rangle = |\psi_i\rangle$.

To see this, consider a state $|\psi\rangle$ written in terms of orthonormal basis-states $|\alpha_j\rangle$, such that $|\psi\rangle = \sum_j c_j|\alpha_j\rangle$ and we would like to write it in a different orthonormal basis $|\beta_j\rangle$. Using the the closure relation, we can write

$$\begin{aligned}|\psi\rangle &= \sum_j c_j|\alpha_j\rangle = \sum_j c_j \mathbb{1}|\alpha_j\rangle = \sum_j c_j \left(\sum_k |\beta_k\rangle\langle\beta_k|\right)|\alpha_j\rangle = \sum_{j,k}\langle\beta_k|\alpha_j\rangle c_j|\beta_k\rangle \\ &= \sum_{j,k} U_{k,j}c_j|\beta_k\rangle\end{aligned} \tag{1.64}$$

The scalars $U_{k,j} = \langle\beta_k|\alpha_j\rangle$ are the matrix elements of a unitary operator U. This is the unitary operator which effects the basis change, $|\alpha_j\rangle = U|\beta_j\rangle$. To prove that this is indeed a unitary operator we need to prove that $U^\dagger U = \mathbb{1}$. The matrix elements of $U^\dagger$ are the transpose conjugate of U, i.e. $(U^\dagger)_{j,k} = U^*_{k,j} = \langle\alpha_j|\beta_k\rangle$.

Thus the elements of $U^\dagger U$ are

$$\begin{aligned}(U^\dagger U)_{j,m} &= \sum_k (U^\dagger)_{j,k}U_{k,m} = \sum_k \langle\alpha_j|\beta_k\rangle\langle\beta_k|\alpha_m\rangle \\ &= \langle\alpha_j|\left(\sum_k |\beta_k\rangle\langle\beta_k|\right)|\alpha_m\rangle = \langle\alpha_j|\mathbb{1}|\alpha_m\rangle \\ &= \delta_{j,m}\end{aligned} \tag{1.65}$$

You can verify that the matrix elements of the identity operator do indeed satisfy $\mathbb{1}_{j,m} = \delta_{j,m}$ since it has 1s on its diagonal and 0 entries everywhere else.

The unitary operators are similar to (in fact are a generalisation of) the rotation operators in two or three dimensional space. The rotation operators indeed share all the above properties, for example, three orthogonal vectors, forming a basis, will remain orthogonal after any rotation. Similarly, rotating a vector does not change its length. For this reason, unitary operators are sometimes referred to as unitary rotations.

Note that it is not conventional to use a "hat" to write unitary operators.

1.19 Projector

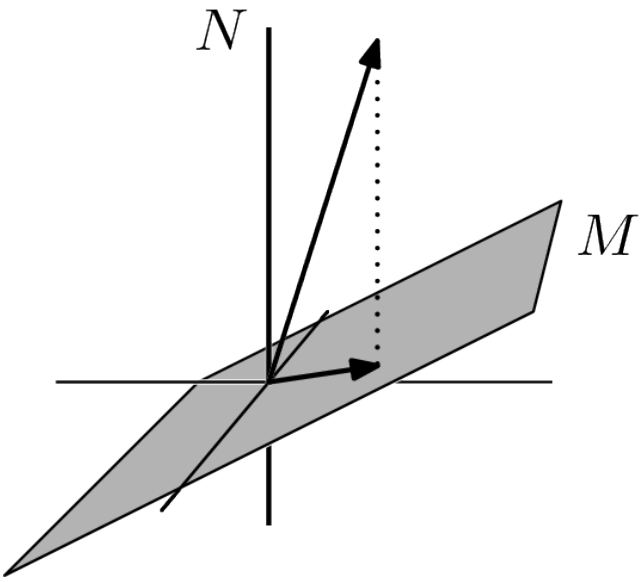

Figure 1.3: This illustration shows the action of a projector projecting a vector in $\mathcal{R}^3$ onto a 2-D plane. Image reproduced from `http://commons.wikimedia.org/wiki/File:Linalg_projection_onto_plane.png`.

Projectors play an important role in quantum measurement.

Definition

The definition of projector is very simple; a projector is any linear operator P which satisfies:

$$P^2 = P \tag{1.66}$$

In other words, a projector is an operator for which multiple application has the same effect as a single application.

We can construct a projector from the product of any (normalised) state $|\psi\rangle$ and its dual $\langle\psi|$.

$$P_{|\psi\rangle} = |\psi\rangle\langle\psi| \tag{1.67}$$

Due to the normalisation condition $\langle\psi|\psi\rangle = 1$ this operator satisfies equation (1.66).

$$P^2_{|\psi\rangle} = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = P_{|\psi\rangle} \tag{1.68}$$

Note again the important difference between $\langle\psi|\psi\rangle$ and $|\psi\rangle\langle\psi|$. One of the strengths of Dirac notation is that the difference between these two orderings is so apparent and the two can be easily distinguished just by their shape on the written page.

More generally, projectors also include sums of operators of this form:

$$P = \sum_{j=1}^{r} |\phi_j\rangle\langle\phi_j| \tag{1.69}$$

where the states $|\phi_j\rangle$ are orthonormal. Any projector satisfying equation (1.66) is called an orthogonal projector. The number r of terms in the decomposition (1.69) is known as the **projector rank**.

The action of a projector is to "project" a vector onto a subspace, by "removing" the components of the vector which cause it to lie outside the sub-space. For example, consider the following matrix operator on $\mathcal{R}^3$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{1.70}$$

You can verify that this matrix is a projector. The action of this operator is to project a general vector

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{1.71}$$

onto its component in the $x - y$ plane,

$$\begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \tag{1.72}$$

You can verify that the $x - y$ plane (or indeed any plane through the origin) forms a subspace.

The *rank* of a projector corresponds to the dimension of the subspace onto which it is projecting.

1.20 Hermitian Operator

Definition

An operator A is called Hermitian (or self-adjoint) if it is equal to its Hermitian conjugate.

$$A^\dagger = A. \tag{1.73}$$

Spectral decomposition

One of the most important properties of Hermitian operators is the spectral decomposition theorem. It states that any Hermitian operator H can be written in the form

$$H = \sum_{j=1}^{d} \lambda_j |\psi_j\rangle\langle\psi_j| \tag{1.74}$$

where d is the dimension of the space on which H acts, the orthonormal states $|\psi_j\rangle$ are *eigenstates* of the operator, and the scalars λ_j are the corresponding *eigenvalues*.

Equation (1.74) is called the *spectral decomposition* of the operator since *spectrum* of the operator is another word for the set of eigenvalues. The theorem represents the well-known fact that Hermitian operators can be diagonalised. Since $\{|\psi_i\rangle\}$ form an orthonormal basis, there exists a unitary U which maps that basis onto another basis set $|e_j\rangle$, i.e. $|\psi_j\rangle = U|e_j\rangle$.

We can rewrite equation (1.74) in terms of the basis $|e_j\rangle$

$$\begin{aligned} H =& \sum_{j=1}^{d} \lambda_j U|e_j\rangle\langle e_j|U^\dagger \\ =& U\left(\sum_{j=1}^{d} \lambda_j |e_j\rangle\langle e_j|\right) U^\dagger \\ =& UDU^\dagger \end{aligned} \tag{1.75}$$

You may be more familiar with the $UDU^\dagger$ form of the spectral decomposition. If we associate the basis vectors $|e_j\rangle$ with the standard basis vectors for $\mathcal{C}^d$, $\vec{e}_j$ introduced above in equation (1.44), then the matrix representation of operator D is a diagonal matrix. For this reason, finding the eigenvalues and eigenstates of an observable is often called *diagonalisation*.

Degenerate eigenvalues

Sometimes, two or more of the eigenvalues in the spectral decomposition, equation (1.74) will coincide. We call these eigenvalues degenerate.

For example consider the following spectral decomposition in a three-dimensional space.

$$H = \lambda_1|\phi_1\rangle\langle\phi_1| + \lambda_2|\phi_2\rangle\langle\phi_2| + \lambda_2|\phi_3\rangle\langle\phi_3| \tag{1.76}$$

The eigenvalue λ_2 is degenerate as it is associated with both $|\phi_2\rangle$ and $|\phi_3\rangle$. We can thus re-write H as follows

$$\begin{aligned} H &= \lambda_1|\phi_1\rangle\langle\phi_1| + \lambda_2\left(|\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|\right) \\ &= \lambda_1 P_1 + \lambda_2 P_2 \end{aligned} \tag{1.77}$$

where $P_1 = |\phi_1\rangle\langle\phi_1|$ and $P_2 = |\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|$ are projectors. The action of projector P_2 projects the state into the sub-space spanned by $|\phi_2\rangle$ and $|\phi_3\rangle$.

We say that this sub-space is the *eigenspace* associated with the degenerate eigenvalue λ_2. Note that any vector in this sub-space is an eigenvalue of H with eigenvalue λ_2 since

$$H\left(\alpha|\phi_2\rangle + \beta|\phi_3\rangle\right) = \alpha\lambda_2|\phi_2\rangle + \beta\lambda_2|\phi_3\rangle = \lambda_2(\alpha|\phi_2\rangle + \beta|\phi_3\rangle) \tag{1.78}$$

In general, the spectral decomposition equation (1.74) can always be written in terms of projectors

$$H = \sum_j \lambda_j P_j \tag{1.79}$$

where the projector P_j projects onto the eigenstate associated with λ_j for non-degenerate eigenvalues or the eigenspace associated with degenerate eigenvalues λ_j.

1.21 Postulate 2: Time-evolution

Now we have covered sufficient mathematical background to look in more detail at Postulates 2 and 3. First, we consider Postulate 2 which describes *time-evolution* in quantum systems.

Postulate 2: Evolution

The evolution of an isolated system is described by the (time-dependent) Schrödinger equation,

$$i\hbar\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle \tag{1.80}$$

where H is the Hamiltonian operator, a **Hermitian operator** which specifies the full dynamics of the system.

Previously, you have learnt a number of methods for solving the Schrödinger equation in different scenarios, for example, by finding the eigenstates of a time-independent Hamiltonian. In this section, we shall take a closer look at the general mathematical description of quantum time evolution. This will provide a base from which we can develop methods for approximating the time-evolution for systems, where the Schrödinger equation cannot be solved exactly.

Time evolution and unitarity

Regardless of the Hamiltonian, there are certain properties which all solutions to the time-dependent Schrödinger equation satisfy.

One of these is unitarity. Evolution under the Schrödinger equation is always *unitary*, which means the state $|\psi(t_1)\rangle$ at time t_1 evolves to state $|\psi(t_2)\rangle$ at t_2 via a unitary *evolution operator*,

$$|\psi(t_2)\rangle = U(t_2, t_1)|\psi(t_1)\rangle. \tag{1.81}$$

This has a number of important consequences, due to the properties of unitary operators we encountered above. For example, unitary transformations preserve norms, so any normalised state will remain normalised under unitary evolution. Also, unitary evolution preserves the inner product between pairs of vectors, so, e.g. sets of orthogonal states remain orthogonal under Schrödinger evolution. We shall study unitary time-evolution in detail in chapter 4 of this course.

1.22 Postulate 3: Measurement

The final postulate of quantum mechanics defines quantum measurement. Measurement is fundamentally important in any theory since it describes the process

by which we learn about a system in an experiment. Measurement in quantum mechanics is strikingly different to measurement in the classical world. In quantum mechanics, measurement outcomes are usually intrinsically probabilistic, and the act of measurement changes the quantum state. These aspects are summarised in Postulate 3.

Postulate 3: Measurement (for discrete systems)

> Associated with every observable quantity is a **Hermitian operator** M with a **spectral decomposition** $M = \sum_j \lambda_j P_j$. The **eigenvalues** λ_j label the possible outcomes of the measurement. When the measurement is performed on a system in state $|\psi\rangle$, the probability that eigenvalue λ_j is returned is $\langle\psi|P_j|\psi\rangle$.
>
> After a measurement which has returned eigenvalue λ_j the state of the system becomes
>
> $$|\psi'\rangle = \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}. \tag{1.82}$$

This postulate tells us that every observable quantity that we measure, whether angular momentum, spin, energy, etc. is represented by a Hermitian operator and its eigenvalues tell us the values of that quantity which can be measured. This postulate assumes that the measured value is discrete. For continuous variables such as position and momentum, it must be modified. We shall present a modified definition for continuous variables in part 4.

The probability of outcome λ_j and the transformation of the state are both determined by projector P_j.

Example 1 - Pauli matrices and the spin-half particle

The state of a spin half particle inhabits a two-dimensional vector space. We can construct matrix representation of the spin-operator in x-, y- and z (and other) directions. These matrices are closely related to a very important family of matrices, the *Pauli matrices*.

Pauli Matrices

The Pauli matrices are the following 2×2 matrices.

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1.83}$$

The Pauli matrices have historically played an important role in many fields of Physics, and have recently gained extra prominence in the field of Quantum Information, where the two-dimensional Hilbert space represents the qubit.

Since they are so important, we will describe their properties in some detail. You can easily verify that the following is true for all $k = x, y, z$

- The matrices are self-inverse $\sigma_k^2 = \mathbb{1}$.
- The matrices are both *unitary* and *Hermitian* $\sigma_k^\dagger = \sigma_k$, $\sigma_k \sigma_k^\dagger = \mathbb{1}$.
- The Pauli matrices *anti-commute*, which means that for $j \neq k$, $\sigma_j \sigma_k = -\sigma_k \sigma_j$.

Spin-half operator matrix representations

The Pauli matrices provide a compact representation for the spin operators for a spin-1/2 particle. We usually represent such operators using the basis defined by the eigenstates of the $\hat{S}_z$ operator, $|\uparrow\rangle$, with eigenvalue $+\hbar/2$, and $|\downarrow\rangle$ with eigenvalue $-\hbar/2$.

In the following representation,

$$|\uparrow\rangle \cong \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |\downarrow\rangle \cong \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{1.84}$$

the matrix representation of $\hat{S}_z$ will be

$$\hat{S}_z \cong \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \frac{\hbar}{2} \sigma_z \tag{1.85}$$

We find that similarly, the other spin operators, in x and y directions can be compactly expressed in terms of Pauli matrices

$$\hat{S}_x \cong \frac{\hbar}{2} \sigma_x \qquad \hat{S}_y \cong \frac{\hbar}{2} \sigma_y \tag{1.86}$$

Note that since the Pauli operators are 2×2 matrices, we can readily diagonalise them, and find their eigenvectors and eigenvalues. For example, the normalised eigenvectors of the Pauli σ_x are $|x+\rangle = (1/\sqrt{2})(|\uparrow\rangle + |\downarrow\rangle)$ for eigenvalue

$+1$ and $|x-\rangle = (1/\sqrt{2})(|\uparrow\rangle - |\downarrow\rangle)$ for -1. Similarly, the normalised eigenstates of σ_y are $|y\pm\rangle(1/\sqrt{2})(|\uparrow\rangle \pm i|\downarrow\rangle)$ for eigenvalues ± 1 respectively.

Spin measurement

Let us now consider the measurement of a spin-half particle in the x-direction. This is a non-degenerate measurement with two outcomes $+\hbar/2$ and $-\hbar/2$. The Hermitian operator which describes it can be written in the spectral decomposition as:

$$\hat{S}_x = \frac{\hbar}{2}|x_+\rangle\langle x_+| + \frac{-\hbar}{2}|x_-\rangle\langle x_-| \tag{1.87}$$

The $+\hbar/2$ outcome is associated with projector $|x_+\rangle\langle x_+|$. Consider a general spin 1/2 state $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$. We can use postulate 3 to calculate the probability of achieving a $+\hbar/2$ outcome in an x-spin measurement.

$$\langle\psi||x_+\rangle\langle x_+||\psi\rangle = |\langle x_+|\psi\rangle|^2 = \frac{1}{2}|\alpha + \beta|^2 \tag{1.88}$$

After the measurement the state is transformed to

$$\frac{|x_+\rangle\langle x_+|\psi\rangle}{\sqrt{\langle\psi|x_+\rangle\langle x_+|\psi\rangle}} = |x_+\rangle \tag{1.89}$$

In your previous courses you may have learnt that the state of the system, after a measurement, corresponds to the eigenstate associated with the measured eigenvalue. Equation (1.89) is compatible with this, so you may wonder why we need a complicated expression such as equation (1.141) to tell us the post-measurement state. As we shall see next, however, the more general form is necessary when one considers a degenerate measurement.

Example 2: a degenerate measurement

Let us consider an example of degenerate measurement using the Hermitian operator with degenerate eigenvalues introduced above in equation (1.76)

$$H = \lambda_1|\phi_1\rangle\langle\phi_1| + \lambda_2\left(|\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|\right) \tag{1.90}$$

Consider a general state $|\psi\rangle = \alpha|\phi_1\rangle + \beta|\phi_2\rangle + \gamma|\phi_3\rangle$ normalised so that $|\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1$.

The projector corresponding to outcome λ_2 is $|\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|$, thus according to postulate 1, the probability of this outcome is

$$\langle\psi|\left(|\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|\right)|\psi\rangle = \langle\psi|\phi_2\rangle\langle\phi_2|\psi\rangle + \langle\psi|\phi_3\rangle\langle\phi_3|\psi\rangle = |\beta|^2 + |\gamma|^2 \tag{1.91}$$

The state is transformed to

$$\begin{aligned} |\psi'\rangle &= \frac{\left(|\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|\right)|\psi\rangle}{\sqrt{\langle\psi|\left(|\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|\right)|\psi\rangle}} \\ &= \frac{\beta|\phi_2\rangle + \gamma|\phi_3\rangle}{\sqrt{|\beta|^2 + |\gamma|^2}} \end{aligned} \tag{1.92}$$

Expectation value

An important quantity in quantum mechanics is the *expectation value*. The expectation value represents the average value of a measurement outcome of observable $\hat{X}$ and is written $\langle\hat{X}\rangle$.

In quantum mechanics, the expectation value for observable $\hat{X}$ given state $|\psi\rangle$ has a simple form;

$$\langle\hat{X}\rangle = \langle\psi|\hat{X}|\psi\rangle \tag{1.93}$$

To see that the expectation value represents the average of the measurement outcomes we use the spectral decomposition of X,

$$\langle\hat{X}\rangle = \langle\psi|(\sum_j \lambda_j|\phi_j\rangle\langle\phi_j|)|\psi\rangle = \sum_j \lambda_j p_j \tag{1.94}$$

where λ_j is the eigenstate of $\hat{X}$ corresponding to the eigenstate $|\phi_j\rangle$ and $p_j = |\langle\phi_j|\psi\rangle|^2$ is the probability of that measurement outcome.

Note that the probabilities which arise in Postulate 3 are themselves expectation values of the projectors, e.g. $\langle\psi|P_j|\psi\rangle = \langle P_j\rangle$. It is therefore often said that expectation values represent all the experimental predictions of quantum mechanics.

Global Phase

Two states which differ only by a global phase, e.g. $|\psi\rangle$ and $|\phi\rangle = e^{i\theta}|\psi\rangle$ have no measurable difference in quantum mechanics. The reason for this is that all expectation values on those states are the same,

$$\langle\phi|\hat{X}|\phi\rangle = e^{-i\theta}e^{i\theta}\langle\psi|\hat{X}|\psi\rangle = \langle\psi|\hat{X}|\psi\rangle. \tag{1.95}$$

For this reason, we are free to arbitrarily choose the global phase in our state vector description, without changing any properties of the state. However, this does not mean that we can ignore the relative phase of terms in a superposition, e.g. $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$ is *not* equivalent to $\alpha|\uparrow\rangle + e^{i\phi}\beta|\downarrow\rangle$[1].

[1]Except for the trivial case where ϕ is a multiple of 2π.

1.23 Density matrices

What is a state?

Postulate one gave us a mathematical definition of a quantum state, in particular:

> A complete description of the system is given by its **state vector**, $|\psi\rangle$, a **unit-normalised vector** in the state space.

By *complete description*, we mean that the state allows us to make the *best* prediction of the outcome of *any* experiment. We have seen above that the outcome of measurements in quantum mechanics can be intrinsically random. This means that there is no way *in principle* to predict with certainty the outcome of measurements. The best prediction we can make is to calculate *probabilities* of measurements and *expectation values*. Since the probability of any outcome can be expressed as an expectation value, we can provide a convenient working definition of a state.

> A **state** is a mathematical object associated with a physical system which allows us to calculate any expectation value - and hence predict the statistics of any measurement on the system.

It is worth emphasising that since expectation values or probabilities are statistical properties. We can only measure them in the lab by repeating our experiment many times, under *identical conditions*.

The state vector description is a powerful one, but it is not the most general. There are some quantum experiments for which no single state vector can give a complete description. These are experiments which have additional randomness or uncertainty, which might mean that either state $|\psi_1\rangle$ *or* $|\psi_2\rangle$ is prepared.

This situation will typically occur due to imperfections in an experimental device, or due to correlation (entanglement) between systems. We shall introduce entanglement in section 1.24.

In principle, one can calculate everything in quantum mechanics using state vectors (as postulate 1 tells us). However, we shall see that it is much more convenient to introduce a new formalism for these cases. This is called the **density matrix** or **density operator** formalism.

Adding classical uncertainty - averages of averages

We have seen that measurements in quantum mechanics usually are intrinsically non-deterministic. However, there is another kind of randomness which can arise in experiments. "Classical randomness" describes the probabilistic behaviour we encountered in classical physics, for example, the outcome of a coin toss or a dice roll. Classical randomness arises, not due to any intrinsic non-determinism (in classical physics all evolution is deterministic) but due to our lack of detailed information about a system. We can characterise the behaviour of a system in terms of probabilities. For example, we say that the probability of a fair coin toss returning heads is $1/2$.

When we consider quantum systems, classical randomness can still occur, for example, in imperfect experimental devices (which only behave as desired with a certain probablility) or in a gas in thermal equilibrium, where the individual properties of any particular atom in a gas are not known, but their probabilities (for example, the speed distribution) can be calculated directly from the thermodynamic properties of the gas.

We therefore need a formalism, which can combine both types of randomness. Remembering that any prediction of quantum mechanics can be expressed as an expectation value, it is worth considering first how classical expectation values (average quantities) can be combined probabilistically. Here's a simple example, imagine that the average temperature on sunny days is 20° but on rainy days it is 10°. If a typical day has a probability p_s of being sunny and $p_r = 1 - p_s$ of being rainy, what is the average temperature of a typical day? To find the overall average temperature we take an *average of the averages* – in this case $(p_s 20 + p_r 10)^\circ$.

If classical randomness influences the preparation of a quantum system, the expectation values must be combined in the same way. Let us illustrate this first with a toy example, and then move to more physical examples below. Consider a lab experiment, where atoms are prepared in one of two possible states $|\psi_1\rangle$ or $|\psi_2\rangle$. You are told that the probabilities of the machine preparing $|\psi_1\rangle$ or $|\psi_2\rangle$ are p_1 and $p_2 = 1 - p_1$ respectively. You will measure observable $\hat{O}$. Given what you know about probabilities, what will be the expectation value for $\hat{O}$?

We know that for atoms prepared in $|\psi_1\rangle$, the expectation value would be

$$\langle \hat{O} \rangle = \langle \psi_1 | \hat{O} | \psi_1 \rangle \tag{1.96}$$

Similarly for atoms in state $|\psi_2\rangle$ the expectation value is

$$\langle \hat{O} \rangle = \langle \psi_2 | \hat{O} | \psi_2 \rangle \tag{1.97}$$

For the case where $|\psi_1\rangle$ is prepared with probability p_1 and $|\psi_2\rangle$ is prepared with probability $p_2 = 1 - p_1$ we must use the average-of-averages rule. This gives the following expression for the expectation value:

$$\langle \hat{O} \rangle = p_1 \langle \psi_1 | \hat{O} | \psi_1 \rangle + p_2 \langle \psi_2 | \hat{O} | \psi_2 \rangle. \tag{1.98}$$

A more compact form - introducing the density operator

Equation (1.98) gives the correct expectation value, but it has a rather complicated form. To calculate it we need to specify both state vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ and their probabilities. It would be more convenient if we could find a single object from which the expectation value (indeed from which *any* expectation value) could be calculated. This object is called a *density operator* or *density matrix* (the two terms are used interchangeably).

To introduce the density operators, we first need one extra linear algebraic tool - the *Trace* of an operator.

Trace

The trace of a matrix is simply the sum of the diagonal elements. For an operator in an inner product space, the trace of that operator is the trace of a matrix representation of that operator. We can write it as follows:

$$\mathrm{Tr}(A) = \sum_j A_{jj} = \sum_j \langle \phi_j | A | \phi_j \rangle \tag{1.99}$$

where $|\phi_j\rangle$ are an orthonormal basis.

From this expression it would appear that the trace of A depends on the orthogonal basis used to define the matrix representation. As we shall see below, however, the trace of an operator is independent of the matrix representation used to compute it, and equation (1.99) is a valid definition with respect to any orthonormal basis.

The trace has a number of useful general properties, which we will quickly summarise.

Linearity of Trace

The trace is linear in its argument, i.e.

$$\text{Tr}(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \text{Tr}(A_1) + \lambda_2 \text{Tr}(A_2) \tag{1.100}$$

This is a direct consequence of its definition, as a linear sum of matrix elements.

Cyclic invariance of trace

The trace is cyclic invariant. That means given the trace of a product of operators A, B and C, the operators may be permuted cyclically, $\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB)$. The proof of this is straightforward, we simply express the matrix multiplication of the matrix representations of A, B and C element-wise, and see that we can reorder the multiplications cyclically

$$\text{Tr}(ABC) = \sum_{j,k,l} A_{jk} B_{kl} C_{lj} = \sum_{j,k,l} B_{kl} C_{lj} A_{jk} = \sum_{j,k,l} C_{lj} A_{jk} B_{kl} \tag{1.101}$$

Basis independence of trace

The definition of the trace of an operator depends upon the particular basis chosen to define the matrix representation of the operator. Here, we shall show that the trace is, in fact, basis independent. Let us consider two orthonormal bases $|\phi_j\rangle$ and $|\psi_j\rangle = U|\phi_j\rangle$. Defined with respect to the first basis, the trace of A is

$$\text{Tr}_\phi(A) = \sum_j \langle \phi_j | A | \phi_j \rangle \tag{1.102}$$

where the suffix ϕ indicates the basis with respect to which the trace is defined.

Defined with respect to the second basis, the trace of A is

$$\text{Tr}_\psi(A) = \sum_j \langle \psi_j | A | \psi_j \rangle \tag{1.103}$$

which we can rewrite, since $|\psi_j\rangle = U|\phi_j\rangle$

$$\begin{aligned} \text{Tr}_\psi(A) &= \sum_j \langle \phi_j | U^\dagger \, A \, U | \phi_j \rangle \\ &= \text{Tr}_\phi(U^\dagger A U) \end{aligned} \tag{1.104}$$

We now use the cyclic invariance of trace.

$$\mathrm{Tr}_\phi(U^\dagger AU) = \mathrm{Tr}_\phi(AUU^\dagger) = \mathrm{Tr}_\phi(A) \tag{1.105}$$

and hence

$$\mathrm{Tr}_\psi(A) = \mathrm{Tr}_\phi(A) \tag{1.106}$$

Since our choice of bases were completely general, this implies that the trace of O is independent of the basis chosen for the matrix representation. We therefore do not need to specify the basis used for the trace, and write Tr without a suffix.

Deriving the density operator

Now let us return to equation (1.98). We want to transform this equation, so that there is a single mathematical object from which the expectation value can be calculated. We will do this using the properties of trace, as follows:

$$\begin{aligned}\langle\hat{O}\rangle &= p_1\langle\psi_1|\hat{O}|\psi_1\rangle + p_2\langle\psi_2|\hat{O}|\psi_2\rangle \\ &= \mathrm{Tr}\left(p_1\langle\psi_1|\hat{O}|\psi_1\rangle\right) + \mathrm{Tr}\left(p_2\langle\psi_2|\hat{O}|\psi_2\rangle\right)\end{aligned} \tag{1.107}$$

It seems odd to insert the trace here, but it is a useful trick. Every scalar quantity such as $p_1\langle\psi_1|\hat{O}|\psi_1\rangle$ is equal to its trace, since a scalar is a 1×1 matrix. We can now use the cyclic invariance and linearity of trace to transform the expression further:

$$\begin{aligned}&\mathrm{Tr}\left(p_1\langle\psi_1|\hat{O}|\psi_1\rangle\right) + \mathrm{Tr}\left(p_2\langle\psi_2|\hat{O}|\psi_2\rangle\right) \\ =&p_1\mathrm{Tr}\left(\hat{O}|\psi_1\rangle\langle\psi_1|\right) + p_2\mathrm{Tr}\left(\hat{O}|\psi_2\rangle\langle\psi_2|\right) \\ =&\mathrm{Tr}\left(\hat{O}\left(p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2|\right)\right) \\ =&\mathrm{Tr}\left(\hat{O}\rho\right)\end{aligned} \tag{1.108}$$

where $\rho = p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2|$ is the *density operator*. The density operator contains all the information needed to calculate any expectation value for the experiment.

In the density operator formalism, the expectation value has a very simple expression in terms of ρ.

$$\langle\hat{O}\rangle = \mathrm{Tr}\left(\hat{O}\rho\right) \tag{1.109}$$

We can generalise this. Assume now that the experiment prepares one of n different pure states $|\psi_j\rangle$, each with probability p_j, such that $\sum_{j=1}^{n} p_j = 1$. Here, as above, the density matrix is simply the sum of the projectors for each pure state, weighted by their respective probability.

$$\rho = \sum_{j=1}^{n} p_j |\psi_j\rangle\langle\psi_j| \tag{1.110}$$

and the expectation value for operator $\hat{O}$ remains $\langle\hat{O}\rangle = \text{Tr}\left(\hat{O}\rho\right)$. Note that states $|\psi_j\rangle$ do *not* need to be orthogonal.

The density operator represents the state of the system

When there is statistical uncertainty in an experiment, the density operator is our best description of the system. We therefore consider it to be *the state* of the system.

You may be worried that ρ is considered the state when "really" the state should be $|\psi_1\rangle$ or $|\psi_2\rangle$. The way to understand this is that the expectation value describes an average over many measurements. In the experiment above, only a proportion of the measurements would be on $|\psi_1\rangle$ and only a proportion on $|\psi_2\rangle$. Since ρ efficiently captures the averaging between these states. For this reason, some textbooks refer to a density matrix as *only* having a meaning when applied to an ensemble of systems. This view is incorrect. As we shall see below, there *are* cases where a single system has to have a density matrix interpretation.

You can get yourself into a philosophical headache about issues like these. When in doubt, it useful to consider the quantum state to represent your "state-of-knowledge" about a system. From this perspective there is no contradiction between assigning ρ as the state or $|\psi_1\rangle$ or $|\psi_2\rangle$ depending on the knowledge you have. It is no different from assigning a different average daily temperature depending on whether you know a day is rainy or dry, or if you have not yet looked out the window to see the weather.

Pure vs mixed states

We have now seen two different types of states, states which *can* be described by a state vector $|\psi\rangle$ alone, and states for which a density matrix is needed. To distinguish between them, we use the terms *pure states* and *mixed states*.

Pure states

A pure state is any state whose density matrix consists of a single projector

$$\rho = |\psi\rangle\langle\psi| \tag{1.111}$$

All calculations for pure states can be made in either the density matrix picture or the state vector picture. The pure states include all the states you encountered before you learned about density operators.

Mixed states

A mixed state is any state which is not a pure state, i.e. its density matrix cannot be written as a single projector. There is no way to represent mixed states using the state vector formalism alone.

Properties of the density operator

A valid density operator is any operator which can be constructed as the sum of a set of projectors

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j| \tag{1.112}$$

It is a linear operator and therefore can be represented as a matrix in any suitable orthonormal basis. For example, with respect to orthonormal basis vectors $|\phi_j\rangle$, we can write

$$\rho = \sum_{j,k} \rho_{j,k} |\phi_j\rangle\langle\phi_k| \tag{1.113}$$

where $\rho_{j,k} = \langle\phi_j|\rho|\phi_k\rangle$ are the matrix elements of the matrix representation.

From the definition in equation (1.112) we can derive that every density operator has the following properties:

- ρ is a *Hermitian* linear operator.

$$\rho = \rho^\dagger \tag{1.114}$$

- ρ has unit trace, which ensures that the probabilities for outcomes of any measurement sum to one.

$$\mathrm{Tr}(\rho) = 1 \tag{1.115}$$

- ρ is a positive operator, which means that all eigenvalues of $\hat{O}$ are non-negative. This ensures that no negative probabilities arise for any measurement.

$$\rho \geq 0 \tag{1.116}$$

Examples

The example above seemed rather unphysical. Why would we expect an experiment to create states according to a probability distribution? And why would we know these probabilities? However, there are *many* practical situations where mixed quantum states arise and the density operator formalism must be used. Here are two examples.

Example 1: Imperfect photon source

A photon is a single particle of light. When a photon exists, we write its state $|1\rangle$ (for 1-photon). We write the state of the vacuum (when no photon is present) $|0\rangle$. Photons are immensely small packets of energy and it is a significant technological challenge to create them. If we could create single photons on demand, they would have practical applications in quantum cryptography and quantum computing. The best single photon sources at present will typically only succeed a proportion of the time, in other words with a probability p they create a photon, but with probability $1 - p$ they fail. Since the success of the photon generation is not indicated by the photon source, we do not know which of the outcomes has occurred. The state of the system produced by the photon source is therefore a mixed state:

$$\rho = p|1\rangle\langle 1| + (1-p)|0\rangle\langle 0| \tag{1.117}$$

Example 2: Systems in Thermal Equilibrium

Consider a system with energy eigenstates $|\psi_n\rangle$, each with (non-degenerate) energy E_n. Statistical mechanics tells us that, in thermal equilibrium at temperature T, the probability that the system is $|\psi_n\rangle$ is given by the Boltzmann factor

$\exp[-\beta E_n]/\sum_m \exp[-\beta E_m]$, where $\beta = k_b/T$. At non-zero temperatures, the system is therefore in a mixed state described by

$$\rho = \frac{\sum_n \exp[-\beta E_n]|\psi_n\rangle\langle\psi_n|}{\sum_m \exp[-\beta E_m]} \tag{1.118}$$

1.24 Compound systems

We have now introduced the most important techniques for representing single systems with discrete observables. In nature, however, physical systems are constantly interacting with the systems around them. In this section, we shall learn how to describe interacting quantum systems.

You will learn how to describe compound systems within the vector space framework, by the introduction of the **Tensor Product**. You will see that this simple mathematical structure has some radical consequences for quantum physics, leading to *entangled states*, the states which are, in some senses, the most non-classical of all. Finally, you shall see how we can construct density matrices for compound systems and that we need density matrix formalism to describe single systems which are entangled with others.

The Tensor product

Consider two systems A and B. Let us assume that we have built up a quantum mechanical description of them, identifying a state space for each, which we will write $\mathcal{A}$ and $\mathcal{B}$. Let us also assume that we have identified an orthonormal basis for these systems, which we shall write $|a_1\rangle, |a_2\rangle, \ldots$ and $|b_1\rangle, |b_2\rangle, \ldots$ respectively.

We would like to find a joint description of these two systems. Furthermore, we would like the description of the joint system to satisfy Postulate 1, i.e. be itself a vector space.

It is clear that this state space needs to contain states which represent, for example, "state $|a_1\rangle$ on system A and state $|b_1\rangle$ on system B" and also "state $|a_2\rangle$ on system A and state $|b_2\rangle$ on system B", and in addition should satisfy the axioms of a vector space. In particular, it needs to be closed under vector addition and scalar multiplication. Thus, one would need to be able to have states which represent "state $|a_1\rangle$ on system A and state $|b_1\rangle$ on system B" + "state $|a_2\rangle$ on system A and state $|b_2\rangle$ on system B". In other words, the principle of superposition must apply to the joint system.

The **tensor product** is the formal name for how we combine Hilbert spaces such that these requirements are satisfied. The symbol for tensor product is $\otimes$ and we write the joint space as $\mathcal{A} \otimes \mathcal{B}$.

Given d_a-dimensional space $\mathcal{A}$ with basis states $\{|a_j\rangle\}$ for $j = 1, 2, \ldots, d_a$ and d_b dimensional space $\mathcal{B}$ with basis states $\{|b_k\rangle\}$ for $k = 1, 2, \ldots, d_b$ we can construct the space $\mathcal{A} \otimes \mathcal{B}$ by identifying a set of orthogonal basis vectors for the joint state space. Basis vectors of the tensor product states are formed by pairing together basis states from $\mathcal{A}$ and $\mathcal{B}$. For example, we write $|a_1\rangle \otimes |b_1\rangle$ to correspond to the state in $\mathcal{A} \otimes \mathcal{B}$, where system A is in state $|a_1\rangle$ and system B is in state $|b_1\rangle$.

A full basis for $\mathcal{A} \otimes \mathcal{B}$ can be constructed by including all $d_a \times d_b$ pairings of basis vectors from the two spaces. Then an arbitrary vector in $\mathcal{A} \otimes \mathcal{B}$ can be written:

$$|\psi\rangle = \sum_{j=1}^{d_a} \sum_{k=1}^{d_b} c_{j,k} |a_j\rangle \otimes |b_k\rangle \tag{1.119}$$

The dimension of the new space is then $d_a \times d_b$.

It is conventional to drop the $\otimes$ sign when context makes it clear that a tensor product is meant. This is particularly common in atomic physics, where one conventionally writes tensor product components in a single ket, separated by commas, e.g. writing $|a_1, b_1\rangle$ for $|a_1\rangle \otimes |b_1\rangle$, and we shall follow this convention in part 2 of the course on quantum angular momentum. Another common convention is to drop the $\otimes$ sign and write $|a_1\rangle|b_1\rangle$.

The component spaces $\mathcal{A}$ and $\mathcal{B}$ form subspaces of the joint tensor product space $\mathcal{A} \otimes \mathcal{B}$.

The tensor product is associative and distributive. This means we can write the tensor product of multiple systems without brackets as $\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{C} \otimes \cdots$.

Example: Two spin-half particles

Let us label one spin A and one spin B. The basis states of A are labelled $|\uparrow\rangle_A$ and $|\downarrow\rangle_A$ and those for B are labelled $|\uparrow\rangle_B$ and $|\downarrow\rangle_B$. We form the basis states for the tensor product space from the four basis pairs

$$\begin{aligned} &|\uparrow\rangle_A|\uparrow\rangle_B \\ &|\uparrow\rangle_A|\downarrow\rangle_B \\ &|\downarrow\rangle_A|\uparrow\rangle_B \\ &|\downarrow\rangle_A|\downarrow\rangle_B \end{aligned} \tag{1.120}$$

A general state for the two-spin system is written

$$|\psi\rangle = \alpha_{\uparrow\uparrow}|\uparrow\rangle_A|\uparrow\rangle_B + \alpha_{\uparrow\downarrow}|\uparrow\rangle_A|\downarrow\rangle_B + \alpha_{\downarrow\uparrow}|\downarrow\rangle_A|\uparrow\rangle_B + \alpha_{\downarrow\downarrow}|\downarrow\rangle_A|\downarrow\rangle_B. \tag{1.121}$$

Tensor products of operators

We can also construct tensor products of operators acting on $\mathcal{A}\otimes\mathcal{B}$. Given operator P acting on $\mathcal{A}$ and operator Q acting on $\mathcal{B}$, the tensor product $P \otimes Q$ acts as you would expect:

$$P \otimes Q|\psi\rangle = \sum_{j=1}^{d_a}\sum_{k=1}^{d_b} c_{j,k}\left(P|a_j\rangle\right) \otimes \left(Q|b_k\rangle\right). \tag{1.122}$$

Operators of the form $P \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \cdots$ which act as the identity on all but one subspace are often written with the shorthand $P_A = P \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \cdots$, $P_B = \mathbb{1} \otimes P \otimes \mathbb{1} \otimes \mathbb{1} \cdots$, where the subscript identifies upon which subsystem the operator acts.

The tensor product can be applied between any state spaces, yet is is most commonly applied between spaces corresponding to clearly defined degrees of freedom, for example, spin and spatial degrees of freedom, or the angular momentum of different particles.

Tensor product in matrix representations

We can formulate the tensor product for matrix representations of states and operators. Just as we saw above we can find a matrix representation of the tensor product. We have seen that, given space A of dimension d_A and space B of dimension d_B, the dimension of the tensor product space $\mathcal{A} \otimes \mathcal{B}$ is $d_A d_B$. This means we will be representing states and operators in the tensor product space by $d_A d_B$-element vectors and $d_A d_B \times d_A d_B$ matrices.

The tensor product rule for both matrices and vectors is as follows. Consider $p_1 \times p_2$ matrix P and a $q_1 \times q_2$ matrix Q acting in space $\mathcal{B}$. The tensor product of these two matrices is a $p_1 q_1 \times p_2 q_2$ matrix written

$$P \otimes Q = \begin{bmatrix} P_{11}Q & P_{12}Q & \cdots & P_{1p_2}Q \\ P_{21}Q & P_{22}Q & \cdots & P_{2p_2}Q \\ \vdots & \vdots & \vdots & \vdots \\ P_{p_1 1}Q & P_{p_1 2}Q & \cdots & P_{p_1 p_2}Q \end{bmatrix} \tag{1.123}$$

which is essentially a matrix made up of a $p_1 \times p_2$ grid of copies of matrix Q each multiplied by an element of P.

Example

Here is an example of the tensor product of two vectors:

$$\begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}3\\4\end{bmatrix} = \begin{bmatrix}1 \times \begin{bmatrix}3\\4\end{bmatrix}\\ 2 \times \begin{bmatrix}3\\4\end{bmatrix}\end{bmatrix} = \begin{bmatrix}3\\4\\6\\8\end{bmatrix} \tag{1.124}$$

Example 2: Spin-half particles

Considering two spin-half particles again, if we associate vector $\begin{bmatrix}1\\0\end{bmatrix}$ with $|\uparrow\rangle$ and $\begin{bmatrix}0\\1\end{bmatrix}$ with $|\downarrow\rangle$ applying the tensor product rules for matrices and vectors we find the vector representation of the four basis states of the tensor product

$$\begin{aligned}
|\uparrow\rangle_A|\uparrow\rangle_B &= \begin{bmatrix}1 \times \begin{bmatrix}1\\0\end{bmatrix}\\ 0 \times \begin{bmatrix}1\\0\end{bmatrix}\end{bmatrix} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}\\
|\uparrow\rangle_A|\downarrow\rangle_B &= \begin{bmatrix}1 \times \begin{bmatrix}0\\1\end{bmatrix}\\ 0 \times \begin{bmatrix}0\\1\end{bmatrix}\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}\\
|\downarrow\rangle_A|\uparrow\rangle_B &= \begin{bmatrix}0 \times \begin{bmatrix}1\\0\end{bmatrix}\\ 1 \times \begin{bmatrix}1\\0\end{bmatrix}\end{bmatrix} = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}\\
|\downarrow\rangle_A|\downarrow\rangle_B &= \begin{bmatrix}0 \times \begin{bmatrix}0\\1\end{bmatrix}\\ 1 \times \begin{bmatrix}0\\1\end{bmatrix}\end{bmatrix} = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}
\end{aligned} \tag{1.125}$$

We see that the basis vectors in the tensor product space are yet again basis vectors in the standard form.

Entanglement

Erwin Schrödinger called entanglement the defining feature of quantum mechanics. Its existence is a simple consequence of the tensor product structure of quantum mechanics, but it has far reaching consequences. Entangled states are the states which are considered most non-classical, and are at the root of many of the puzzles and paradoxes we face when we try to interpret quantum mechanics. They also form a key resource for quantum computation and quantum cryptography and are therefore now being studied with much greater scrutiny than before.

We call a state in space $\mathcal{A} \otimes \mathcal{B}$ a *product state* if it can be written as a tensor product. Let us consider two two level systems, e.g. spin 1/2 particles, which we shall label A and B. The following are product states:

$$\begin{aligned} |\psi_1\rangle &= |\uparrow_A\rangle|\uparrow_B\rangle \\ |\psi_2\rangle &= |\downarrow_A\rangle|\uparrow_B\rangle \\ |\psi_3\rangle &= \frac{1}{2}\left(|\uparrow_A\rangle|\uparrow_B\rangle + |\uparrow_A\rangle|\downarrow_B\rangle + |\downarrow_A\rangle|\uparrow_B\rangle + |\downarrow_A\rangle|\downarrow_B\rangle\right) \\ &= \frac{1}{2}(|\uparrow_A\rangle + |\downarrow_A\rangle)(|\uparrow_B\rangle + |\downarrow_B\rangle) \end{aligned} \tag{1.126}$$

Product states are the states we obtain when we use the tensor product to combine the description of two systems, whose state is described by a state vector. It is easy to reverse the process, and read off from the product state, the individual states of the two systems. For example, in state $|\psi_1\rangle$ it is easy to see that the state of system A is $|\uparrow_A\rangle$, and the state for system B is $|\uparrow_B\rangle$. Note that $|\psi_3\rangle$ was not initially written in product form, but remains a product state since it can be factored into tensor product form.

Surprisingly, there exist states which cannot be factored in this way. Remember that, since this is a vector space, we can always form states by superposition (vector addition) of other states. This allows us to construct states which cannot be factored - *entangled states*. An entangled state is a state which is impossible to write in tensor product form, i.e. all non-product states are entangled states.

An example of an entangled state is

$$|\psi_4\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_A\rangle \otimes |\uparrow_B\rangle + |\downarrow_A\rangle \otimes |\downarrow_B\rangle\right) \tag{1.127}$$

Since an entangled state cannot be written as a direct tensor product, there is no way (in the state-vector formalism) to write down a separate "state-vector-for-system-A" and "state-vector-for-system-B". We shall see below that it is possible

to write a "reduced state" of systems A and B as mixed states, in the density matrix formalism, but in doing so, one loses some of the information in the state. The information lost relates to the correlations between the systems.

Correlations in measurement outcomes are a defining feature of entangled states. Consider state $|\psi_4\rangle$. By postulate 3 (or simply by inspection), if both spins are measured along the z-axis, i.e. operators $\hat{S}_z \otimes \mathbb{1}$ and $\mathbb{1} \otimes \hat{S}_z$ are measured, the outcomes are correlated, i.e. either both outcomes are up or both outcomes are down. We can re-express state $|\psi_4\rangle$ in the eigenbasis of $\hat{S}_x$, states $|x+\rangle = (1/\sqrt{2})(|\uparrow\rangle + |\downarrow\rangle)$ and $|x-\rangle = (1/\sqrt{2})(|\uparrow\rangle - |\downarrow\rangle)$ and find

$$|\psi_4\rangle = \frac{1}{\sqrt{2}} \left(|x+_A\rangle \otimes |x+_B\rangle + |x-_A\rangle \otimes |x-_B\rangle \right). \tag{1.128}$$

This implies that the outcome of these measurements is also correlated, i.e. if $|x+\rangle$ is measured on the first spin it will also be measured on the second spin. Note also that the probability of any of these measurements is entirely random. There is a fifty:fifty chance of getting either outcome in either measurement. The measurements exhibit global correlation but local randomness, a feature which is characteristic of entangled states.

Now imagine that the two spins are located very far from one another. After the first one is measured, postulate 3 tells us that the state of the system is "collapsed" to a product state, e.g. in the first example to $|\uparrow_A\rangle|\uparrow_B\rangle$ or $|\downarrow_A\rangle|\downarrow_B\rangle$. This seems to imply that the state of system B is changed the instant that system A is measured, which seems contrary to spirit of special relativity, which tells us that no signal can travel faster than light. To resolve this, let us look at the situation from a physicist making a measurement on system B. Until the signal of system A reaches them, he or she does not know the outcome of A's measurement and their best description of the spin is as a probabilistic ensemble, with probability $1/2$ the state is $|\uparrow\rangle$ and with probability $1/2$ it is $|\downarrow\rangle$. The density matrix which describes this is $(1/2)(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)$. We call this state the *reduced state* of system B. Without receiving the measurement outcome at system A, this description is the best one which our physicist at system B can assign. It is independent of the outcome of the measurement of system A, and hence, from our physicists perspective, the state of his spin does not change at all when system A is measured, and no signal travels instantaneously.

We can run through the same argument for $\hat{S}_x$ measurements, and find that the density matrix for system B, would be $(1/2)(|x+\rangle\langle x+| + |x-\rangle\langle x-|)$, which reexpressed in the $|\uparrow\rangle$, $|\downarrow\rangle$ basis is $(1/2)(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)$. This "reduced state" for system B is the same regardless of whether system A is measured in the x or

z-basis.

This is reassuring, if these states were different, then the physicist at system B could determine which of the measurements at system A had been made, and a signal (the choice of measurement basis) have would be sent instantaneously.

In fact, regardless of the measurement made on system A, one can show that the reduced state of system B is the same. As we shall see below, the reduced state system A can be calculated directly from the pure entangled state describing the joint system of A and B, and does not depend on any action (measurement or other transformation) on system B. Otherwise, it would be possible to use entangled states to signal faster-than-light, in direct contradiction of special and general relativity. Fortunately, at the level of observable quantities, quantum mechanics and special relativity remain compatible with one another.

Entangled states play a central role in quantum computing, and you will learn many more of their properties in Prof Bose's lecture course on Quantum Computing next term. They lead to a number of puzzling consequences (such as the violation of Bell inequalities) which challenge the way we understand quantum mechanics. I will discuss these in my lecture on "Interpreting Quantum Mechanics".

There are far more down-to-earth reasons for studying entangled states, however. Namely that almost all quantum systems which are interacting become entangled, and almost all quantum systems in nature are interacting!

Density Matrices for compound systems

We construct density matrices for compound systems, in precisely the same way as for isolated systems. For example, for systems A and B with orthonormal bases $|a_j\rangle$ and $|b_k\rangle$ respectively, the vector space for the joint system, is described by the orthonormal basis $|a_j\rangle \otimes |b_k\rangle$ which we shall write, $|a_j, b_k\rangle$ for compactness. In the same way as for single systems, compound systems may be described by density matrices. For example, a general density operator ρ on this joint system takes the form

$$\rho = \sum_{j,k} \sum_{l,m} \rho_{j,k,l,m} |a_j, b_k\rangle\langle a_l, b_m| \tag{1.129}$$

where the matrix element $\rho_{j,k,l,m} = \langle a_j, b_k|\rho|a_l, b_m\rangle$.

In compound systems, mixed states may be also exhibit entanglement. A detailed discussion of entanglement in mixed states is beyond the syllabus of this course, and is still an active area of research.

Reduced states for entangled systems

In the previous section, we saw that if two systems A and B are in an entangled state, we can no longer separate the state description to write down individual state vectors for system A and B. Remember that a state is a mathematical object which allows to predict the expectation values for any observable and the probabilities for any measurement outcome. The *reduced state* for system A (or B) is the state which allows one to calculate the value of any observable which acts solely on system A, in other words, for any measurement of the properties of system A which ignores system B.

The reduced state is obtained by an operation called *partial trace*.

Partial Trace

Recall the definition of Trace. For our compound system,

$$\text{Tr}(\rho) = \sum_{j,k} \rho_{j,k,j,k} \tag{1.130}$$

where $\rho_{j,k,l,m} = \langle a_j, b_k|\rho|a_l, b_m\rangle$. Note that the sum ranges over both indices j and k.

The partial trace is an operation which allows us to "delete" a system from our tensor product compound system, to give us a **reduced state** of the remaining system. For example, to delete system A, we perform the partial trace over system A:

$$\text{Tr}_A(\rho) = \sum_{l,m} \left(\sum_j \rho_{j,l,j,m} \right) |b_l\rangle\langle b_m| = \rho_B \tag{1.131}$$

The sum over j indicates the index of the system we are "removing". We say that system A is "traced out" of the state description. The partial trace operation leaves us with a state for system B. It is called a partial trace as we sum over indices for system A alone. The state which remains is constructed from the basis states for B alone. Note the notation. Trace over A is signified by Tr_A, reduced state for system B is signified by ρ_B.

The reduced state ρ_B is the density matrix which describes the state of system B, if it is considered individually. It can be used to calculate the expectation value for every observable which acts solely on B (i.e. every measurement on B alone), via the standard formula

$$\langle \hat{O} \rangle = \text{Tr}(\hat{O}\rho_B) \tag{1.132}$$

$$\mathrm{Tr}_A(\rho) = \sum_{l,m} \left(\sum_j \rho_{j,l,j,m} \right) |b_l\rangle\langle b_m| = \rho_B \tag{1.133}$$

Similarly, the reduced state for system A is obtained by the partial trace over B;

$$\mathrm{Tr}_B(\rho) = \sum_{j,k} \left(\sum_l \rho_{j,l,k,l} \right) |a_j\rangle\langle a_k| = \rho_A \tag{1.134}$$

Reduced states of entangled systems

It is possible to show that the reduced state of a system is a pure state, only if that system has no entanglement with any other physical system. This means that the reduced state of an entangled system is always a **mixed state**. This is a very different kind of mixed state to the one introduced above. In that case, the mixed state description was necessary due to classical randomness - here we see that mixed states also arise due to entanglement.

Sometimes textbooks call a mixture due to classical randomness a "proper mixture" and one due to entanglement "improper". In calculations, (and therefore for all practical purposes) there is *no distinction* between improper and proper mixtures, so I recommend that you do not make a distinction between them.

Example: Spontaneous Decay

The discussion of entangled states and mixed states may seem rather abstract, so here is a concrete example. It is well known that an atom in an excited state $|e\rangle$ can decay spontaneously to a lower energy ground state $|g\rangle$ with the emission of a photon. Spontaneous decay occurs due to the interaction with the electromagnetic vacuum around the atom. The full theory of spontaneous emission is rather involved (the atom interacts with all possible frequencies, polarisations and momenta of the em-field), so I present here a toy model, which captures the essential features of the process. Let us write down the vacuum state of the electromagnetic field as $|0\rangle$ and the state where one photon is present as $|1\rangle$.

We can describe spontaneous emission heuristically via the following time-varying state

$$|\psi(t)\rangle = \exp[-\kappa t]|e\rangle|0\rangle + \sqrt{1-\exp[-2\kappa t]}|g\rangle|1\rangle \tag{1.135}$$

For all times $t > 0$ this state is entangled. We would like to find a state description for the atom alone. To do this, we calculate its reduced state. First we write $|\psi(t)\rangle$ as as density matrix

$$\begin{aligned}|\psi(t)\rangle\langle\psi(t)| =& \exp[-2\kappa t]|e\rangle|0\rangle\langle e|\langle 0| + (1-\exp[-2\kappa t]|g\rangle|1\rangle\langle g|\langle 1| \\ &+ (\exp[-\kappa t])\sqrt{1-\exp[-2\kappa t]}|e\rangle|0\rangle\langle g|\langle 1| \\ &+ (\exp[-\kappa t])\sqrt{1-\exp[-2\kappa t]}|g\rangle|1\rangle\langle e|\langle 0|\end{aligned} \tag{1.136}$$

If we take the partial trace over the photon mode, we recover the following reduced state for the atom

$$\rho_{\text{atom}} = \text{Tr}_{\text{light}}(|\psi(t)\rangle\langle\psi(t)|) = \exp[-2\kappa t]|e\rangle\langle e| + (1-\exp[-2\kappa t])|g\rangle\langle g| \tag{1.137}$$

This mixed state has precisely the probabilistic behaviour we would expect. At time t, there is probability $\exp[-2\kappa t]$ that the system is in the excited state. Thus, the probability of non-emission decays exponentially. The reduced state is sufficient for us to deduce this and calculate any other expectation value for the atom alone, during the decay process.

Note that even though the state of the joint system is pure, for all finite times $t > 0$, the state of the atom is mixed. This is due to the entanglement between the atom and the emitted light.

Unitary evolution of density operators

We have seen that under the Schrödinger equation, a state vector evolves unitarily $|\psi(t)\rangle = U(t)|\psi\rangle$. This implies that a density matrix will evolve as follows.

If at time $t = 0$ the system is in state

$$\rho(0) = \sum_j p_j |\psi_j\rangle\langle\psi_j| \tag{1.138}$$

then it will evolve under unitary evolution operator $U(t)$ as

$$\rho(t) = \sum_j p_j U(t)|\psi_j\rangle\langle\psi_j|U(t)^\dagger = U(t)\rho U(t)^\dagger \tag{1.139}$$

Note that this implies that a state which is *initially pure* will *stay pure* under unitary evolution. In fact there are only two ways in which mixed states can arise in quantum mechanics,

- if there is some classical uncertainty in the experiment - (this corresponds to our lack of knowledge about the fine detailed dynamics of a system)
- or if, during the evolution, a system has become entangled with other system(s). In this case the global state remains pure, but the state of the system itself becomes mixed.

In many systems of interest, the particles we are studying do become entangled with other bodies during their natural evolution. We call such systems *open quantum systems* since they are interacting with systems outside our control (in contrast, systems with no external interactions are called closed). We call these external systems the *environment* of the system we are studying. For example, the environment of an atom might include electro-magnetic fields in its vicinity, as well as other nearby particles.

In open quantum dynamics, due to this interaction, systems are usually described by mixed states, and thus their evolution equations must be expressed in the density operator formalism. You will learn about such evolution in chapter 6 of this course, on open quantum system dynamics.

1.25 Continuous systems

So far, we have focused our systems representing *discrete variables*, i.e. systems where measured quantities only take a discrete set of values. However, we know that some of the most important types of observable in nature (e.g. position and momentum) take a continuous range of values. In this section we shall see how *continuous variable* systems may be represented in Dirac notation, and that, apart from some subtleties which arise due to the continuous nature of values, the structure is essentially the same as for the finite dimensional case.

From wave functions to Dirac notation

The wave function representation of the position of a particle $\psi(x)$ is the first representation of a quantum state which most students encounter, so it is worth seeing how it fits in with the Dirac notation formalism we have been developing.

Recall that for the position of a (one-dimensional) quantum particle represented by wave-function $\psi(x)$, the probability of finding that a particle in the range $a \leq x \leq b$ is

$$\text{Prob}(a \leq x \leq b) = \int_a^b |\psi(x)|^2 dx. \tag{1.140}$$

We will see, in this section, that with every such wave function, we can associate a state vector $|\psi\rangle$ which inhabits an infinite-dimensional state space, and that essentially the same formalism as introduced above, with a few important changes, can be used to calculate properties of continuous systems. There are some important, and subtle mathematical difficulties which arise in infinite-dimensional spaces. In the main are issues regarding the convergence of infinite sums, and the limits of infinite sequences. These matters can be very technical, and we shall *not* discuss them in this course. I recommend the textbooks by Ballentine and Peres, as good starting points for a detailed discussion of these points. For now, and to the level of rigour we shall adopt in this section, you may consider $|\psi\rangle$ to be a vector in an infinite-dimensional complex inner product space.

From discrete to continuous, from sums to integrals

In physics, when we switch from discrete variables to continuous ones, a common feature is that *sums* are replaced by *integrals*. Our transition from discrete

to continuous will reflect this. When considering a measurement of a continuous quantity, it is impossible (even in classical physics) to measure the quantity precisely, a real number might require infinitely many digits of precision in order to be exactly expressed. Our measurements, therefore, always cut the continuous variable into discrete intervals. We see this reflected in equation (1.140), probability in this expression for the measured value to lie in the range from a to b. Using equation (1.140) as a crutch, we shall start by finding a way to express probabilities such as expressed in this equation, in Dirac notation. To do this, we need a continuous version of the quantum mechanical measurement postulate.

1.26 Measurement for continuous variable observables

Recall the measurement postulate for discrete observables.

> Associated with every observable quantity is a **Hermitian operator** M with a **spectral decomposition** $M = \sum_j \lambda_j P_j$. The **eigenvalues** λ_j label the possible outcomes of the measurement. When the measurement is performed on a system in state $|\psi\rangle$, the probability that eigenvalue λ_j is returned is $\langle\psi|P_j|\psi\rangle$.
>
> After a measurement which has returned eigenvalue λ_j the state of the system becomes
>
> $$|\psi'\rangle = \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}. \tag{1.141}$$

The key object here is the observable, $M = \sum_j \lambda_j P_j$ a Hermitian operator which contains the allowed measurement values λ_j and defines the sub-spaces corresponding to states of definite value for this observable. Let us restrict our analysis to non-degenerate measurements. The observable then takes the form $M = \sum_j \lambda_j |\phi_j\rangle\langle\phi_j|$, where $|\phi_j\rangle$ is the eigenstate associated with eigenvalue λ_j. For a generic state $|\psi\rangle$, the probability of an outcome λ_j is $\text{Prob}(\lambda_j) = |\langle\phi_j|\psi\rangle|^2$.

We need a continuous variable version of this postulate. The usual approach, when extrapolating from discrete to continuous variables, is to replace *sums* by *integrals*. Hence for a non-degenerate position observable, one might guess that

the operator would have the form

$$\hat{x} = \int_{-\infty}^{\infty} x|x\rangle\langle x|dx. \tag{1.142}$$

In fact, we will see that equation (1.142) is the correct expression for a position operator in one-dimension. In analogue to the spectral decomposition for discrete operators, the continuous values x, and $|x\rangle$ and $\langle x|$ represent eigenvectors and their duals. We shall study the properties of $|x\rangle$ and $\langle x|$ in more detail below.

First, given a measurement operator of the form of equation (1.142), we shall propose an expression for the probability of a measurement of this $\hat{x}$ operator on a state $|\psi\rangle$. Since measurements of continuous variables can only be made to finite precision we must consider the probability of the position falling in a range $a \leq x \leq b$.

To construct an expression for this, we work again in analogy with the discrete case. There, to calculate the probability of a range of mutually exclusive discrete outcomes (e.g. λ_1, λ_2 or λ_3) we sum their respective probabilities, e.g.

$$\text{Prob}(\lambda_1 \text{ or } \lambda_2 \text{ or } \lambda_3) = \sum_{j=1}^{3} \text{Prob}(\lambda_j) = \sum_{j=1}^{3} |\langle\phi_j|\psi\rangle|^2 \tag{1.143}$$

In the continuous case, we replace the sum by an integral and write the probability of a measurement of $\hat{x}$ to have an outcome within a range $a \leq x \leq b$ to be as follows:

$$\text{Prob}(a \leq x \leq b) = \int_a^b \langle\psi|x\rangle\langle x|\psi\rangle dx = \int_a^b |\langle x|\psi\rangle|^2 dx \tag{1.144}$$

If we compare this with equation (1.140), we see that the two expressions become equal if we make the identification

$$\langle x|\psi\rangle = \psi(x). \tag{1.145}$$

We shall consider this to be a Dirac notation *definition* of the wave-function. For this to make any sense, we first need to look at the properties of $|x\rangle$ in more detail.

1.27 Properties of the $|x\rangle$ states

The vectors $|x\rangle$ which appeared in the above equations, have not yet been described in detail. They represent the eigenstates (eigenvectors) of the operator $\hat{x}$

corresponding to eigenvalue x. To avoid confusing x and $\hat{x}$, when we write about specific eigenstates, we will often use the notation x_0 for the eigenvalue and $|x_0\rangle$ for the state.

Let $|x_0\rangle$ be an eigenstate (eigenvector) of operator $\hat{x}$ with eigenvalue x_0. Thus

$$\hat{x}|x_0\rangle = x_0|x_0\rangle \tag{1.146}$$

It is natural to ask what wave-function is associated with $|x_0\rangle$. Via equation (1.145) this must be:

$$\hat{x}|x_0\rangle = \int_{-\infty}^{\infty} x|x\rangle \underbrace{\langle x|x_0\rangle}_{\text{Wave-function of } |x_0\rangle} dx. \tag{1.147}$$

This equation becomes consistent with equation (1.146) if we allow the wave-function $\langle x|x_0\rangle$ to be a Dirac delta function,

$$\langle x|x_0\rangle = \delta(x - x_0). \tag{1.148}$$

Aside: Dirac delta function

You have encountered the Dirac delta function in previous courses, so we shall merely recap its properties here. It can be defined in many ways:

Definition 1: Inside an integral

$$\int_{-\infty}^{\infty} \delta(x - x_0) f(x) dx = f(x_0), \tag{1.149}$$

which expresses the behaviour of the delta function inside an integral. More intuitive is its definition as a limit.

Definition 2: As a limit

The delta function can be defined as the limit of many functions (essentially any function which can be squeezed to a single infinitely thin, infinitely high peak, while the area under the peak remains unchanged at one.) For example, expressed as a Gaussian,

$$\delta(x - x_0) = \lim_{\epsilon \to 0} \frac{1}{2\sqrt{\pi\epsilon}} e^{-\frac{(x-x')^2}{4\epsilon}}. \tag{1.150}$$

A further useful definition can be derived from the Fourier transform,

Definition 3: From the Fourier transform

$$\delta(x - x_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x-x')}. \tag{1.151}$$

Using any of those definitions, one can derive general properties such as $\int \delta(x)dx = 1$ and $\delta(\lambda x) = \delta(x)/|\lambda|$, where λ is a positive or negative real number.

Properties of the $|x\rangle$ basis

The eigenstates of $\hat{x}$ with eigenvalue x_0, in the range of continuous values between $-\infty$ and $+\infty$, form a basis of this infinite-dimensional Hilbert space. Each state $|x_0\rangle$ has as its wave-function a Dirac-delta function centred at point x_0.

Recall that for discrete observables, eigenstates of non-degenerate observables formed an orthonormal basis $\langle\phi_j|\phi_k\rangle = \delta_{j,k}$. The equivalent expression for position eigenstates $|x_0\rangle$ is equation (1.148). We call such a basis delta-orthonormal.

Closure relation

Similar to bases defined by discrete observables, the states $|x\rangle$ satisfy a closure relation.

$$\int_{-\infty}^{\infty} |x\rangle\langle x| dx = \mathbb{1} \tag{1.152}$$

Any state may thus be expressed in the position basis as follows

$$|\psi\rangle = \mathbb{1}|\psi\rangle = \int_{-\infty}^{\infty} |x\rangle\langle x|\psi\rangle dx = \int_{-\infty}^{\infty} \psi(x)|x\rangle dx \tag{1.153}$$

with the wave-function $\psi(x)$ taking the role of amplitudes in the discrete case (c.f. $|\psi\rangle = \sum_j |\phi_j\rangle$). Thus $\psi(x)$ is sometimes called the position representation of $|\psi\rangle$.

Inner product

Consider two states $|\psi_1\rangle$ and $|\psi_2\rangle$. Their inner product can be obtained by exploiting closure relation (1.152)

$$\begin{aligned}\langle\psi_1|\psi_2\rangle &= \langle\psi_1| \int_{-\infty}^{\infty} |x\rangle\langle x| dx\ |\psi_2\rangle \\ &= \int_{-\infty}^{\infty} \langle\psi_1|x\rangle\langle x|\psi_2\rangle dx \\ &= \int_{-\infty}^{\infty} \psi_1(x)^* \psi_2(x) dx\end{aligned} \tag{1.154}$$

The inner product is therefore represented by the "overlap integral" between wave-functions.

Norm

We shall use the norm defined by the inner product

$$|||\psi\rangle|| = \sqrt{\langle\psi|\psi\rangle} \tag{1.155}$$

which, inserting the closure relation is equal to

$$|||\psi\rangle|| = \sqrt{\int_{-\infty}^{\infty} |\psi(x)|^2 dx} \tag{1.156}$$

A normalised state satisfies

$$\int_{-\infty}^{\infty} |\psi(x)|^2 = 1 \tag{1.157}$$

The position basis states $|x\rangle$ cannot be normalised, since $\langle x|x\rangle = \delta(0)$ is not a number (the central peak of the delta-function can be thought of as tending to infinity). All physical states must be normalisable, otherwise probabilities cannot be calculated. Thus $|x\rangle$ basis states are not physical states, and cannot be prepared (of course one can construct physical states which are super-positions of these states). This fact should be unsurprising, since the $|x\rangle$ represent states with infinite precision in their position, and we know that continuous variables can only be measured up to finite precision.

Such states are sometimes called "improper states". Even though they are unphysical on their own, they form such a useful basis to represent other states, that they are still very important.

1.28 Position representations of operators

We have seen two different kinds of operators in this chapter, operators acting on a vector space

$$\hat{O}|\psi\rangle = |\psi'\rangle \tag{1.158}$$

and differential operators acting on wavefunctions

$$\tilde{O}\psi(x) = \psi''(x) \tag{1.159}$$

In this section, we will see how the two forms of operator are related. To distinguish them, we shall here write differential operators with a tilde and vector space operators with a hat. For example, for position

$$\tilde{x} = x \qquad \hat{x} = \int_{-\infty}^{\infty} |x\rangle\langle x| dx \tag{1.160}$$

$$\tilde{p} = \frac{\hbar}{i}\frac{\partial}{\partial x} \qquad \hat{p} = \int_{-\infty}^{\infty} |p\rangle\langle p| dp \tag{1.161}$$

Consider operator $\hat{O}$ which maps state $|\psi_1\rangle$ onto state $|\psi_2\rangle$, i.e. $|\psi_2\rangle = \hat{O}|\psi_1\rangle$. The associated differential operator is $\tilde{O}$ which acts as follows:

$$\tilde{O}\psi_1(x) = \psi_2(x) \tag{1.162}$$

which we rewrite

$$\tilde{O}\langle x|\psi_1\rangle = \langle x|\psi_2\rangle = \langle x|\hat{O}|\psi_1\rangle \tag{1.163}$$

Hence $\tilde{O}$ and $\hat{O}$ are related via

$$\tilde{O}\langle x|\psi\rangle = \langle x|\hat{O}|\psi\rangle \tag{1.164}$$

for all states $|\psi\rangle$.

For example,

$$\begin{aligned}
\tilde{x}\ \langle x|\psi\rangle &= \langle x| \int_{-\infty}^{\infty} x'|x'\rangle\langle x'|dx'|\psi\rangle \\
&= \int_{-\infty}^{\infty} x'\langle x|x'\rangle\langle x'|\psi\rangle dx' \\
&= \int_{-\infty}^{\infty} x'\delta(x-x')\langle x'|\psi\rangle dx' \\
&= x\ \psi(x)
\end{aligned} \tag{1.165}$$

as required.

The position representation of the momentum operator is $(\hbar/i)\partial/\partial x$. We can define a corresponding operator in our vector space analogously to the position operator

$$\hat{p} = \int_{-\infty}^{\infty} p|p\rangle\langle p|dp \tag{1.166}$$

where

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \tag{1.167}$$

One can verify that the wave-functions associated with states $|p\rangle$ are Dirac delta-normalised plane-waves, which are, as required, eigenfunctions of $\tilde{p}$.

1.29 Typical spatial Hamiltonians

For many physical problems we write Schrödinger's equation in a spatial representation. Kinetic energy is given by the operator form of the non-relativistic classical value $p^2/2m$,

$$\frac{\tilde{p}^2}{2m} = -\hbar^2 \frac{1}{2m} \frac{\partial^2}{\partial x^2} \tag{1.168}$$

The potential energy operator is usually position dependent $\tilde{V}(x)$. The spatial representation of this Hamiltonian has the familiar form:

$$\tilde{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \tilde{V}(\vec{r}). \tag{1.169}$$