Linear Algebra Notes for PHY356

Michael Luke

University of Toronto
Last Modified: April 6, 2021

ABSTRACT: These notes cover some of the linear algebra required for PHY356 at the University of Toronto. They are drawn mostly from Shankar's text. Note that both Shankar and Cohen-Tannoudji et. al. are much more thorough in their discussion, but these notes are an attempt to present the key bits of linear algebra as succinctly as possible – do not attempt to find much rigour in these notes. Please let me know if any typos/errors/etc. - the notes may be updated from time to time (check the date on this page).

Contents

1. Vectors and Vector Spaces

The first postulate of quantum mechanics is that the state of any physical system is represented by a mathematical object called a "state vector", which lives in a particular kind of vector space called a Hilbert space. Furthermore, observables (that is, things we can physically measure) are represented by operators (more precisely, Hermitian operators, which we will define shortly) acting vectors in the Hilbert space. You have presumably studied

vectors and operators in linear algebra, at least in finite dimensional vector spaces, where you represented vectors in n dimensions by column vectors

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \tag{1.1}$$

(where the c_i's are the components of the vector in some basis) and operators by $n \times n$ matrices

$$\begin{pmatrix} O_{11} & O_{12} & \ldots & O_{1n} \\ O_{21} & O_{22} & \ldots & O_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ O_{n1} & O_{n2} & \ldots & O_{nn} \end{pmatrix}. \tag{1.2}$$

The same techniques you learned in linear algebra will be used in this course to do calculations in quantum mechanics, but with the extra twist that in order to discuss quantum mechanical wavefunctions, we will have to generalize all of these concepts to infinite dimensional vector spaces. The generalization will be straightforward, but just to make everything as clear as possible, let's start at the beginning.

1.1 Vector Spaces

In the first few lectures of this course, we will introduce several different objects which we will call vectors. First of all, we have the usual geometrical vectors ("arrows with magnitude and direction") that you are used to. But then we will meet something more abstract called a "state vector" in quantum mechanics. So the first thing we need to do is figure out what exactly we mean by this term "vector." Mathematicians have a very general definition of vectors - these obey a number of properties that geometrical vectors obviously obey, but they are more general. In particular, with the more general definition, a vector doesn't necessarily have magnitude and direction.

To remind you when we are dealing with a more general definition of vectors, we will denote a vector by

$$|V\rangle \tag{1.3}$$

rather than the more familiar notation $\vec{V}$, which we will reserve for geometric vectors which actually point somewhere (like position $\vec{r}$, momentum $\vec{p}$, or angular momentum $\vec{L}$). Then we make the following definition:

Definition: A **linear vector space** $\mathbb{V}$ is a collection of objects $|V\rangle$, $|W\rangle$, ..., called **vectors** for which there exists

1. A rule for forming the sum $|V\rangle + |W\rangle$, and

2. A rule for multiplying a vector by a scalar to get another vector: $|aV\rangle \equiv a|V\rangle$

and which obey a number of natural requirements:

- Closure (the sum of any two vectors is also a vector): $|V\rangle + |W\rangle \in \mathbb{V}$

- Associativity of vector addition: $|V\rangle + (|W\rangle + |Z\rangle) = (|V\rangle + |W\rangle) + |Z\rangle$.

- Commutativity of vector addition: $|V\rangle + |W\rangle = |W\rangle + |V\rangle$

- Distributivity of scalar multiplication: $a(|V\rangle + |W\rangle) = a|V\rangle + a|W\rangle$

- Associativity of scalar multiplication: $a(b|V\rangle) = ab|V\rangle$

In addition, there exists a null vector[1] denoted $|0\rangle$, which satisfies

$$|V\rangle + |0\rangle = |V\rangle \tag{1.4}$$

for any $|V\rangle$, as well as an inverse vector $|-V\rangle$ for every $|V\rangle$, which satisfies

$$|V\rangle + |-V\rangle = |0\rangle. \tag{1.5}$$

That's it. You'll note that these are very general requirements - in particular, there's nothing that says anything about magnitude or direction. Vectors are just a bunch of objects with rules for adding them and rules for multiplying them by a scalar. As we'll see, with this definition, all sorts of things are vectors, including geometrical vectors like you're used to, but also matrices, functions, and quantum mechanical states.

The numbers $a, b, \ldots$ which multiply the vectors in the previous definition are called the **field** over which the vector space is defined. If $a, b, .. \in \mathbb{R}$, then $\mathbb{V}$ is called a **real** vector space, whereas if $a, b, .. \in \mathbb{C}$, then $\mathbb{V}$ is called a **complex** vector space. Geometric vectors are elements of a real vector space; for example, when we write a velocity

$$\vec{v} = v_x \hat{x} + v_y \hat{y}$$

the numbers v_x and v_y are real numbers. In quantum mechanics, state vectors $|\psi\rangle$ are elements of a complex vector space. For example, when we write, for a two-state system in which a spin can be up or down

$$|\psi\rangle = \alpha|\chi_{\text{up}}\rangle + \beta|\chi_{\text{down}}\rangle$$

the numbers α and β are complex.

It's easy to see that geometrical vectors obey the requirements for vectors given above, but so do lots of things that don't look anything like arrows; for example (you should check this!)

1. the set of all 2×2 matrices

2. the set of all functions $f(x)$ defined on the interval $0 \leq x \leq L$.

Of course not every collection of objects that can be added together that you can think of forms a vector space - for example, does the set of all functions $f(x)$ defined on the interval $0 \leq x \leq L$ satisfying $f(0) = 4$ form a vector space? Why not?

---

[1]A word of advance warning: later on in this course we will use the notation $|0\rangle$ to denote the ground state of a system, *not* the null vector. Sorry about that.

1.2 Dimensions and Basis Vectors

For geometrical vectors, we have an intuitive understanding of the dimension of the vector space. For example, when we write

$$\vec{r} = r_x \hat{x} + r_y \hat{y} \tag{1.6}$$

we know we are dealing with a vector in two dimensions; similarly, in three dimension we write

$$\vec{r} = r_x \hat{x} + r_y \hat{y} + r_z \hat{z}. \tag{1.7}$$

What we mean by saying a (geometrical) vector lives in two dimensions is that we can have a maximum of **two** linearly independent vectors, and any other vector can be written as a linear combination of these two. For example in the figure, any other vector in this two dimensional space can be written $\vec{v} = a\vec{v}_1 + b\vec{v}_2$. If we were in three dimensions we'd need to add another vector in the third direction, and then we could write any vector as a linear combination of those three.[2]
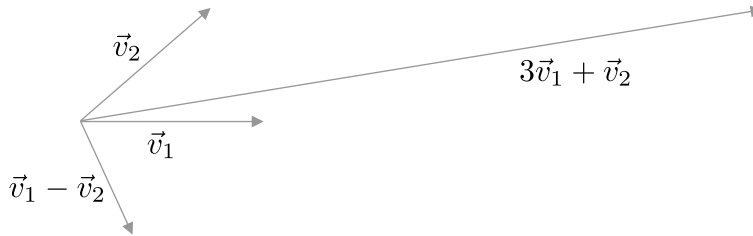
Figure 1: Any vector in two dimensions can be written as a linear combination of $\vec{v}_1$ and $\vec{v}_2$.

We can use the same notion to define the dimension of an arbitrary vector space:

Definition: A vector space has **dimension** n if it can have at most n linearly independent[3] vectors. We can denote the vector space $\mathbb{V}^n(\mathbb{R})$ or $\mathbb{V}^n(\mathbb{C})$, depending on whether it is defined over a real or complex field.

Example: What is the dimension of the space of 2×2 matrices?

Let's define the following vectors:

$$|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, |2\rangle = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, |3\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, |4\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then an arbitrary 2×2 matrix may be written

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a|1\rangle + b|2\rangle + c|3\rangle + d|4\rangle. \tag{1.8}$$

Thus, the space of 2×2 matrices is a four dimensional vector space (either real or complex, depending on whether the coefficients a, b, c, d are real or complex).

---

[2]Note that the vectors v_1 and v_2 don't have to be orthogonal - we haven't even introduced this concept yet!

[3]A set of vectors of linearly independent if no member of the set can be written as a linear combination of all the others.

Example: What is the dimension of the space of functions $f(x)$ defined on $0 \leq x \leq L$?

We will come back to this problem shortly, but for the moment, convince yourself that this is an *infinite* dimensional vector space - that is, no matter how many linearly independent functions $f_i(x)$ I have, I can always write a function $g(x)$ that can't be written as a linear combination of the $f_i(x)$'s. Don't worry if the idea of an infinite dimensional vector space seems daunting at this stage; we'll get a lot of practice with them in this course. One of the important results of these notes is that essentially all of the linear algebra that you are used to for finite dimensional vector spaces can be carried over to infinite dimensional vector spaces, or spaces of functions. These functions will be the familiar wavefunctions of quantum mechanics you studied in second year.

Now, notice in the first of these two examples that we could express an arbitrary vector (in this case, a 2×2 matrix) as a linear combination of four linearly independent vectors $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle\}$; we call these *basis vectors*. Again, this is a completely general concept: in an n-dimensional vector space, we can choose n linearly independent vectors $|i\rangle$ as basis vectors. We can then write any vector in the space as a linear combination

$$|V\rangle = \sum_{i=1}^{n} v_i |i\rangle \tag{1.9}$$

where the $|i\rangle$'s are basis vectors, and the numbers v_i are called the *components* of the vector. (Thus, in Eq. (1.7), $\hat{x}$, $\hat{y}$ and $\hat{z}$ are the basis vectors, and r_x, r_y and r_z are the components of $\vec{r}$.) It's then easy to show that to add vectors, you just need to add their components:

$$|V\rangle + |W\rangle = \sum_i (v_i + w_i)|i\rangle \tag{1.10}$$

so the i'th component of the vector $|V\rangle + |W\rangle$ is $v_i + w_i$.

1.3 Inner Product Spaces

As we have already said, a vector space does not need to have any notion of magnitude and direction associated with it. But recall that for geometrical vectors we can define the dot product of two vectors $\vec{a}$ and $\vec{b}$ by $\vec{a} \cdot \vec{b} = |a||b|\cos\theta$, where θ is the angle between the two vectors. Clearly the dot product tells us about the magnitudes ($|a|$, $|b|$) and directions ($\cos\theta$) of the two vectors. We can define something analogous on any vector space, called the **inner product** (sometimes called the scalar product, or dot product). The inner product of two vectors $|V\rangle$ and $|W\rangle$ is denoted

$$\langle V|W\rangle$$

and satisfies the following conditions:

1. $\langle V|W\rangle = \langle W|V\rangle^*$

2. $\langle V|V\rangle \geq 0$, and $\langle V|V\rangle = 0$ iff $|V\rangle = |0\rangle$

3. $\langle V|(a|W\rangle + b|Z\rangle) = a\langle V|W\rangle + b\langle V|Z\rangle$

where as usual the $*$ denotes complex conjugation. A vector space with an inner product defined is called an *inner product space*. The second requirement allows us to define the length or magnitude of a general vector

$$|V| \equiv \sqrt{\langle V|V\rangle} \tag{1.11}$$

in the same way that $|a| = \sqrt{\vec{a} \cdot \vec{a}}$ for a geometrical vector, and only the null vector has zero magnitude. The first requirement may look new, since you are used to dealing with real vector spaces, but it guarantees that $\langle V|V\rangle = \langle V|V\rangle^*$, so that the magnitude of a vector is real, even in a complex vector space. But it also means you have to be careful when evaluating inner products of sums of vectors! From the general definition of a vector space, we know that the vector $|aW + bZ\rangle = a|W\rangle + b|Z\rangle$, so

$$\langle V|aW + bZ\rangle = a\langle V|W\rangle + b\langle V|Z\rangle. \tag{1.12}$$

However, we also have

$$\langle aW + bZ|V\rangle = \langle V|aW + bZ\rangle^* = a^*\langle V|W\rangle^* + b^*\langle V|Z\rangle^* = a^*\langle W|V\rangle + b^*\langle Z|V\rangle. \tag{1.13}$$

Eq. (1.12) shows that the inner product is linear with respect to the second vector, but Eq. (1.13) shows that it is "anti-linear" with respect to the first vector.

We say that geometrical vectors of nonzero length are orthogonal if their dot product is zero (so $\cos\theta = 0$); now that we've defined the inner product, we can talk about general vectors being orthogonal if their inner product vanishes, even if they have no geometrical interpretation. In particular, it is usually convenient to choose the basis vectors in our vector space to be orthogonal, and to have unit norm (or length); that is, if $|i\rangle$ and $|j\rangle$ denote any two basis vectors, then

$$\langle i|j\rangle = \delta_{ij} \equiv \begin{cases} 1, \text{if } i = j \\ 0, \text{if } i \neq j \end{cases}. \tag{1.14}$$

The symbol δ_{ij} is called the Kronecker delta. You've probably seen this in other courses. If not, you should get very familiar with it as we'll be using it a lot.

We can use Eq. (1.14) to express the inner product of two vectors in terms of their components:

$$\langle V|W\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i^* w_j \langle i|j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i^* w_j \delta_{ij} = \sum_{i=1}^{n} v_i^* w_i \tag{1.15}$$

(where we have used the properties of the Kronecker delta to turn the double sum into a single sum, since all terms with $j \neq i$ vanish in the second sum). This just generalizes the usual formula for the dot product of geometrical vectors in terms of their components,

$$\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + a_z b_z \tag{1.16}$$

to arbitrary vectors in real or complex vector spaces. Note that for a complex vector space, you have to be careful about complex conjugating the components of the vector on the left.

1.4 Bra's and Ket's

This notation $|V\rangle$ for vectors and $\langle V|W\rangle$ for inner products is due to Dirac, and is known as Dirac Notation. We will use it exclusively in this course. The objects $|V\rangle$ we are using to describe vectors are known in Dirac's jargon as "kets" (the reason for this will become clear in a moment). But we saw when introducing the inner product that we have sneakily introduced another object, denoted $\langle W|$ - we will call this a "bra". Thus, when we write an inner product

$$\langle W|V\rangle \tag{1.17}$$

we combine these two object together into a "bra(c)ket", hence the jargon. But what is this new object $\langle W|$ that we've introduced? Well, for any vector (or ket) $|V\rangle$, there is a corresponding associated bra $\langle V|$, so we would expect (correctly) that the objects $\langle V|$ also form a vector space. Furthermore, this can't be the vector space $\mathbb{V}$ since we can't add bra's and ket's: $|V\rangle + \langle W|$ isn't something we've defined, and so it doesn't make any sense; bra's and ket's must therefore live in different vectors spaces.[4] Indeed, formally, the objects $\langle V|$ belong to something called a dual space to the vector space $\mathbb{V}$, but we don't need to worry about the details of dual spaces here (Cohen-Tannoudji et. al discuss them in Chapter II.B.2, and Shankar has a briefer discussion in Chapter 1.3). For our purposes, we can just think of the bra $\langle V|$ as something like the complex conjugate (actually the Hermitian conjugate, which we will define shortly) of the ket $|V\rangle$.

We can see this more clearly by introducing one more piece of notation. It is often convenient to represent orthogonal basis vectors as column vectors; that is, for each basis vector in an n-dimensional vector space, we define an associated n-component column vector. Labelling the basis vectors $|1\rangle$, $|2\rangle, ...|n\rangle$, we write[5]

$$|1\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, |2\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, |n\rangle \leftrightarrow \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{1.18}$$

Then we can represent any vector $|V\rangle$ by a column vector containing its components in this basis:

$$|V\rangle = \sum_i v_i|i\rangle \leftrightarrow \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \tag{1.19}$$

---

[4]That means if you ever end up with an expression in which you are adding bra's and ket's together, **you have made a mistake**. Like all good notation, it is designed to stop you from doing silly things.

[5]Note that I don't use an equals sign here, but rather a double-headed arrow. The basis vectors are not equal to the column vectors, but the column vectors provide one way to represent the basis vectors. The double-headed arrow therefore means "is represented as, in a particular basis." As we will see later, it will sometimes be convenient to switch bases so that the vectors $|1\rangle$, $|2\rangle$, etc. are represented by different column vectors.

Similarly, each basis vector $|i\rangle$ has a corresponding bra, $\langle i|$, with which we can associate a row vector:

$$\langle 1| \leftrightarrow \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}, \langle 2| \leftrightarrow \begin{pmatrix} 0 & 1 & \dots & 0 \end{pmatrix}, \dots, \langle n| \leftrightarrow \begin{pmatrix} 0 & 0 & \dots & 1 \end{pmatrix}. \tag{1.20}$$

Now, if $|W\rangle = \sum_i w_i |i\rangle$, then from Eq. (1.13) we must have

$$\langle W| = \sum_i w_i^* \langle i| \tag{1.21}$$

and so we can represent any bra $\langle W|$ by a row vector whose components are the *complex conjugate* of the components of $|W\rangle$:

$$\langle W| \leftrightarrow \begin{pmatrix} w_1^* & w_2^* \dots w_n^* \end{pmatrix}. \tag{1.22}$$

$\langle W|$ is therefore represented by the *adjoint* of $|W\rangle$:

$$\begin{pmatrix} w_1^* & w_2^* \dots w_n^* \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}^\dagger \tag{1.23}$$

where the adjoint of a matrix (a column vector being a matrix with one column and n rows) is defined as

$$\Omega_{ij}^\dagger = \Omega_{ji}^* \tag{1.24}$$

The inner product of two vectors is then just given by matrix multiplication:

$$\langle W|V\rangle = \begin{pmatrix} w_1^* & w_2^* & \dots & w_n^* \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = w_1^* v_1 + w_2^* v_2 + \dots + w_n^* v_n \tag{1.25}$$

in agreement with Eq. (1.15).

Note that in an orthonormal basis, you can find the i'th component of the vector $|V\rangle$ by taking the inner product of $|V\rangle$ with $|i\rangle$:

$$\langle i|V\rangle = \langle i| \sum_j v_j |j\rangle = \sum_j v_j \langle i|j\rangle = v_i. \tag{1.26}$$

2. Linear Operators

2.1 Definition and Examples

An *operator* Ω is an object which acts on a vector and gives another vector:

$$\Omega|V\rangle = |V'\rangle \tag{2.1}$$

They also act on bra's:

$$\langle V|\Omega = \langle V''| \tag{2.2}$$

(note that for a general operator, as we will see in a moment, we do not necessarily have $V'' = V'$). *Linear* operators, which we will be mostly interested in, obey

$$\Omega\left[\alpha|V\rangle + \beta|W\rangle\right] = \alpha\Omega|V\rangle + \beta\Omega|W\rangle. \tag{2.3}$$

We will generally use the terms "linear operator" and "matrix" interchangeably; that's because *in a given basis*, we can always write the action of a linear operator on a vector as a matrix multiplying a column vector. Let's show this: in Eq. (2.1) we can write $|V\rangle = \sum_i v_i|i\rangle$ (and similarly for $|V'\rangle$). We then have

$$v_i' = \langle i|V'\rangle = \langle i|\Omega|V\rangle = \langle i|\Omega\sum_j v_j|j\rangle = \sum_j v_j\langle i|\Omega|j\rangle \equiv \sum_j \Omega_{ij}v_j \tag{2.4}$$

where in the last equality we have defined the *matrix element* of Ω between the states $|i\rangle$ and $|j\rangle$,

$$\Omega_{ij} \equiv \langle i|\Omega|j\rangle. \tag{2.5}$$

We call this a matrix element because the final equation, which relates the components of $|V'\rangle$ to those of $|V\rangle$,

$$v_i' = \sum_j \Omega_{ij}v_j \tag{2.6}$$

just corresponds to the i'th row of the matrix multiplication

$$\begin{pmatrix} v_1' \\ v_2' \\ \vdots \\ v_n' \end{pmatrix} = \begin{pmatrix} \Omega_{11} & \Omega_{12} & \dots & \Omega_{1n} \\ \Omega_{21} & \Omega_{22} & \dots & \Omega_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{n1} & \Omega_{n2} & \dots & \Omega_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \tag{2.7}$$

Thus, just as we can represent vectors in an n-dimensional vector space by n-component column vectors, we can represent linear operators acting on that space by $n \times n$ matrices. Note that in this example the matrix element is taken between two basis states; in general, we will use the term matrix element to refer to any object of the form $\langle\psi_1|\Omega|\psi_2\rangle$ for arbitrary states $|\psi_1\rangle$ and $|\psi_2\rangle$.

Example: The Identity Operator

The identity operator $\mathbb{1}$ is the simplest operator we can defined on a vector space. It simply leaves all vectors unchanged when it acts on them:

$$\mathbb{1}|V\rangle = |V\rangle \tag{2.8}$$

Its matrix elements are simple to work out:

$$\mathbb{1}_{ij} = \langle i|\mathbb{1}|j\rangle = \langle i|j\rangle = \delta_{ij} \tag{2.9}$$

or in matrix form,

$$\mathbb{1} \leftrightarrow \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \tag{2.10}$$

so $\mathbb{1}$ is represented in any orthonormal basis by the unit matrix.

Example: Projection Operators

Projection operators will show up frequently in our discussions of quantum mechanics. They are analogous to operators on geometrical vectors which project out their components in different directions. The projection operator onto the state $|i\rangle$ is defined by

$$\mathbb{P}_i \equiv |i\rangle\langle i|. \tag{2.11}$$

That may look odd at first glance - the projection operator is made up of a bra and a ket, but they're in the wrong order to be an inner product.[6]. When we write a bra and a ket in this way we don't get an inner product, but an operator. We can see this by acting on a vector:

$$\mathbb{P}_i|V\rangle = |i\rangle\langle i|V\rangle = v_i|i\rangle \tag{2.12}$$

$\mathbb{P}_i$ therefore acts on a vector and gives another vector, which means that it is indeed an operator. Furthermore, it returns a vector pointing along the $|i\rangle$ direction with magnitude v_i - that is, it returns the projection of $|V\rangle$ along the direction $|i\rangle$. To keep our notation consistent, let's use Dirac notation to represent geometrical vectors: we will denote the unit vector in the x direction by $|\hat{x}\rangle$ and the unit vector in the y direction by $|\hat{y}\rangle$. If $|V\rangle$ were a geometrical vector in two dimensions, we could define projection operators $\mathbb{P}_x = |\hat{x}\rangle\langle\hat{x}|$ and $\mathbb{P}_y = |\hat{y}\rangle\langle\hat{y}|$ onto the x and y axes, as shown in Fig. 2.[7]
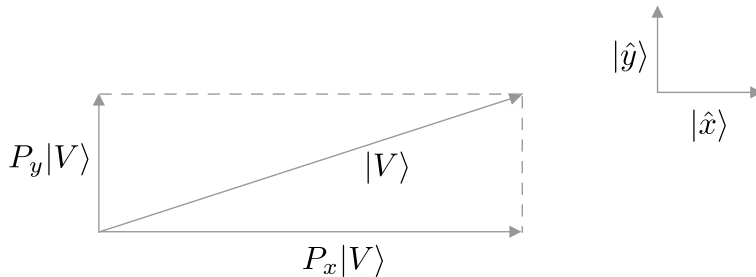
Figure 2: A vector V and its projections $P_x\vec{V}$ and $P_y\vec{V}$ onto the $\hat{x}$ and $\hat{y}$ basis vectors.

In two dimensions, if you project a vector onto the $\hat{x}$ axis, projecting it again onto the same axis doesn't change it. In other words, $\mathbb{P}_x^2|V\rangle = \mathbb{P}_x|V\rangle$, or more generally $\mathbb{P}_x^2 = \mathbb{P}_x$ (and similarly for $\mathbb{P}_y$). If you project a vector onto the $\hat{x}$ axis and then project it onto

---

[6]In Dirac notation, order matters! $|V\rangle\langle W| \neq \langle W|V\rangle$!

[7]In the usual geometrical notation, we would write $\mathbb{P}_x\vec{V} = \hat{x}(\hat{x} \cdot \vec{V})$, so we would write $\mathbb{P}_x = \hat{x}\hat{x}\cdot$, which is somewhat awkward notation. Dirac notation is proving its superiority here.

the $\hat{y}$ axis, you get zero, so $\mathbb{P}_x\mathbb{P}_y = 0$. Finally, if you add the projection of the vector onto the $\hat{x}$ axis with the projection onto the $\hat{y}$ axis, you get the vector back: $\mathbb{P}_x + \mathbb{P}_y = \mathbb{1}$. These relations easily generalize to projection operators in any vector space. The first two properties are easily derived:

$$\mathbb{P}_i\mathbb{P}_j = |i\rangle\langle i|j\rangle\langle j| = \delta_{ij}|i\rangle\langle j| = \delta_{ij}|i\rangle\langle i| \tag{2.13}$$

so

$$\mathbb{P}_i\mathbb{P}_j = \delta_{ij}\mathbb{P}_i. \tag{2.14}$$

(Note that since the Kronecker delta sets the expression to zero unless $i = j$, we can simply take $i = j$ on the projection operator on the right-hand side of Eq. (2.13).) This compactly gives the two relations $\mathbb{P}_i^2 = \mathbb{P}_i$ and $\mathbb{P}_i\mathbb{P}_j = 0$ if $i \neq j$. The third relation comes from

$$|V\rangle = \sum_i v_i|i\rangle = \sum_i \langle i|V\rangle|i\rangle = \sum_i |i\rangle\langle i|V\rangle = \sum_i \mathbb{P}_i|V\rangle \tag{2.15}$$

which immediate gives the *completeness* relation

$$\sum_i \mathbb{P}_i = \mathbb{1}. \tag{2.16}$$

The completeness relation may look trivial, but we will use it over and over again in this course. For example, we can use it to prove that the matrix representing the product of two operators is just the matrix product of the two matrices. Consider two linear operators, Ω and Λ, and the product $\Omega\Lambda$. Then we can cleverly insert the identity matrix in the following derivation

$$(\Omega\Lambda)_{ij} = \langle i|\Omega\Lambda|j\rangle = \langle i|\Omega\mathbb{1}\Lambda|j\rangle$$
$$= \sum_k \langle i|\Omega|k\rangle\langle k|\Lambda|j\rangle = \sum_k \Omega_{ik}\Lambda_{kj}$$

and the last expression is just the matrix product of Ω and Λ. We'll find many more uses for the completeness relation as we go on.

2.2 Adjoints

Note that if we have a ket $|\alpha V\rangle$, where α is a complex number, then $|\alpha V\rangle = \alpha|V\rangle$ (from our original definition of a vector space). But then the corresponding bra is $\langle \alpha V| = \langle V|\alpha^*$, because of the complex conjugation in the inner product. Similarly, if we act an operator on a state

$$|V'\rangle = \Omega|V\rangle \tag{2.17}$$

we can define the *adjoint* $\Omega^\dagger$ of the operator Ω:

$$\langle V'| \equiv \langle V|\Omega^\dagger. \tag{2.18}$$

It's an easy exercise to show that the matrix elements of Ω and $\Omega^\dagger$ are related:

$$\Omega_{ij}^\dagger = \Omega_{ji}^* \tag{2.19}$$

which is just the definition (1.24) of the adjoint of a matrix you will have seen in a linear algebra course. Another useful property which is also easy to prove (try it) is

$$(\Omega\Lambda)^\dagger = \Lambda^\dagger \Omega^\dagger \tag{2.20}$$

so to take the adjoint of the product of two operators you take the adjoint of the two operators and reverse the order of multiplication.

2.2.1 Hermitian Conjugation in Dirac Notation

A ket $|\psi\rangle$ and its corresponding bra $\langle\psi|$ are said to be "Hermitian conjugates" of one another. Similarly, because of the relations (2.17) and (2.18), we say that $A^\dagger$ is the Hermitian conjugate of A – thus, we will often use the terms "Hermitian conjugate" and "adjoint" interchangeably. Since a number is a one by one matrix, we can also say that the Hermitian conjugate of a constant is its complex conjugate.

To find the Hermitian conjugate of the inner product of two vectors, we note that from Section (1.3),

$$\langle\psi_1|\psi_2\rangle^\dagger = \langle\psi_1|\psi_2\rangle^* = \langle\psi_2|\psi_1\rangle \tag{2.21}$$

and so to take the Hermitian conjugate of an inner product we replace bra's with the corresponding ket's (and vice versa) *and* reverse the order of the factors. Similarly, when taking the Hermitian conjugate of a product of operators, we reverse the order (see Eq. (2.20)) in addition to taking the Hermitian conjugates of the original operators. We can generalize this to come up with rules to take the Hermitian conjugate of any expression consisting of bra's, ket's, operators, and constants (see CT, pg. 120). To obtain the Hermitian conjugate/adjoint of any expression,

- Replace

  - the constants by their complex conjugates
  - the ket's by the corresponding bra's, and vice versa
  - the operators by their adjoints

- Reverse the order of the factors (although the position of the constants doesn't matter – this includes inner products and matrix elements.)

For example, $\Omega = \lambda\langle u|AB|v\rangle|w\rangle\langle\psi|$ is an operator (make sure you understand why – λ and $\langle u|AB|v\rangle$ are numbers). Its adjoint is $\Omega^\dagger = |\psi\rangle\langle w|\langle v|B^\dagger A^\dagger|u\rangle\lambda^* = \lambda^*\langle v|B^\dagger A^\dagger|u\rangle|\psi\rangle\langle w|$. Similarly, $|\psi\rangle = \lambda|u\rangle\langle v|w\rangle$ is a ket. Its Hermitian conjugate is $\langle\psi| = \langle w|v\rangle\langle u|\lambda^* = \lambda^*\langle w|v\rangle\langle u|$.

2.3 Unitary and Hermitian Operators

There are two types of linear operators we will be particularly interested in in quantum mechanics: Hermitian operators and Unitary operators. Let's consider these in turn.

2.3.1 Unitary Operators

Unitary operators obey the relation

$$UU^\dagger = U^\dagger U = \mathbb{1}. \tag{2.22}$$

Thus, the adjoint of a a unitary operator is the inverse of the operator, $U^\dagger = U^{-1}$. Unitary operators are important in quantum mechanics because they preserve the norm of a state vector: If $\langle \psi | \psi \rangle = 1$, and $|\psi'\rangle = U|\psi\rangle$, then $\langle \psi' | \psi' \rangle = \langle \psi | U^\dagger U | \psi \rangle = \langle \psi | \psi \rangle = 1$ as well.

Changing Bases: One important use of unitary operators in quantum mechanics is to effect a change of basis.

Just as a vector is a geometrical object independent of basis, the definition of a linear operator in Eqs. (2.1–2.3) makes no reference to any particular basis. When we want to represent a linear operator as a matrix, as in Eq. (2.5), however, the representation depends on the basis we are working in. Frequently in quantum mechanics we will know the form of an operator in one basis, but then need to know it form in another basis. In this section we will show how to do this.

Let's suppose we have an N-dimensional vector space with one set of orthonormal basis vectors which we denote $\{|i\rangle\}$ and another which we denote $\{|\omega_i\rangle\}$, $i = 1..N$. Let us define an operator U which acts on the basis vector $|i\rangle$ and returns $|\omega_i\rangle$ for $i = 1..N$:

$$U|i\rangle = |\omega_i\rangle, i = 1..N. \tag{2.23}$$

Since both sets of basis vectors are normalized, U is unitary:

$$\langle \omega_i | \omega_i \rangle = \langle i | U^\dagger U | i \rangle = 1 \Rightarrow U^\dagger U = \mathbb{1}. \tag{2.24}$$

It is straightforward to work the form of the matrix representing U in the $\{|i\rangle\}$ basis:

$$U_{ij} = \langle i | U | j \rangle = \langle i | \omega_j \rangle. \tag{2.25}$$

In other words, U_{ij} is an $N \times N$ unitary matrix whose j'th column consists of the components of $|\omega_j\rangle$ in the $\{|i\rangle\}$ basis.

Now let's suppose we have some operator Λ, and we know its matrix elements in the $\{|i\rangle\}$ basis:

$$\Lambda_{ij} = \langle i | \Lambda | j \rangle. \tag{2.26}$$

We would now like to write the components of Λ in the $\{|\omega_i\rangle\}$ basis; denote this matrix by $\tilde{\Lambda}_{ij}$, where

$$\tilde{\Lambda}_{ij} = \langle \omega_i | \Lambda | \omega_j \rangle. \tag{2.27}$$

From Eq. (2.23), this is

$$\tilde{\Lambda}_{ij} = \langle i | U^\dagger \Lambda U | j \rangle = (U^\dagger \Lambda U)_{ij} \tag{2.28}$$

so we can express Λ in a new basis just by multiplying the matrix representing it on the left by $U^\dagger$ and on the right by U.

Example: Consider a two-dimensional vector space spanned by two basis vectors $|1\rangle$ and $|2\rangle$; that is, in the $(|1\rangle, |2\rangle)$ basis the vector $|\psi\rangle = \alpha|1\rangle + \beta|2\rangle$ is represented by the column vector

$$|\psi\rangle \leftrightarrow \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{2.29}$$

Let us define another set of orthonormal basis vectors

$$|\omega_1\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle), \quad |\omega_2\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle) \tag{2.30}$$

You can easily check that in the $(|1\rangle, |2\rangle)$ basis the operator U is represented by the matrix

$$U_{ij} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{2.31}$$

Now suppose we have an operator which is represented by the matrix

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{2.32}$$

in the $(|1\rangle, |2\rangle)$ basis. In the $(|\omega_1\rangle, |\omega_2\rangle)$ basis, it will have the form

$$\tilde{\sigma}_y = U^\dagger \sigma_y U = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \tag{2.33}$$

Similarly if we have the operator

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.34}$$

in the $(|1\rangle, |2\rangle)$ basis, it will have the form

$$\tilde{\sigma}_x = U^\dagger \sigma_x U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.35}$$

in the $(|\omega_1\rangle, |\omega_2\rangle)$ basis, which immediately tells us that $|\omega_1\rangle$ and $|\omega_2\rangle$ are eigenvectors of σ_x with eigenvalues 1 and -1, respectively.

2.3.2 Hermitian Operators

Hermitian operators are equal to their own adjoint:

$$\Omega^\dagger = \Omega \tag{2.36}$$

(this, as we will see, is the operator analogue of a quantity being real). Hermitian operators are important in quantum mechanics because, as we discuss in class, observable quantities correspond to Hermitian operators. In particular, for any observable quantity Ω, the only allowed results of a measurement are eigenvalues of Ω, and after measuring an eigenvalue of Ω the system will then collapse into the corresponding eigenstate[8] . Two important properties of Hermitian operators (which you should be able to prove) are

---

[8]Recall that if a vector $|\varphi_i\rangle$ is an eigenvalue of an operator Ω, it obeys the equation $\Omega|\varphi_i\rangle = \omega_i|\varphi_i\rangle$ where ω_i is the eigenvector of $|\varphi_i\rangle$ (note that Ω is an operator, but ω_i is a number).

1. Eigenvalues of a Hermitian operator are real, and

2. Eigenvectors of a Hermitian matrix corresponding to different eigenvalues are orthogonal.

These are important physically because the first property ensures that measured quantities are real, not complex, numbers, whereas the second ensures that you can always choose an orthonormal set of basis vectors consisting of the eigenvectors of an observable.

There is an important subtlety for the second property, however. If an operator has degenerate eigenvectors – that is, two eigenvectors with the same eigenvalue – any linear combination of the two eigenvectors is also an eigenvector with the same eigenvalue: if $\Omega|\varphi_1\rangle = \omega|\varphi_1\rangle$ and $\Omega|\varphi_2\rangle = \omega|\varphi_2\rangle$, then $\Omega(a|\varphi_1\rangle + b|\varphi_2\rangle) = \omega(a|\varphi_1\rangle + b|\varphi_2\rangle)$. Therefore, any vector in the two dimensional subspace spanned by $|\varphi_1\rangle$ and $|\varphi_2\rangle$ is an eigenvector of Ω with the same eigenvalue. Since there are now an infinite number of eigenvectors with the same eigenvalue, they can't all be orthogonal.

This isn't a problem, though - since we have a two dimensional subspace of eigenvectors with the same eigenvalue, we can always find two orthogonal eigenstates in that subspace. Hence, even in the degenerate case, we can always choose all of the eigenvectors to be orthogonal.

Example: Consider the Hermitian operator with matrix representation

$$\Omega = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}. \tag{2.37}$$

You should also recall from a linear algebra course that the eigenvalues ω_i are solutions to the characteristic equation

$$\det(\Omega - \omega) = 0 \tag{2.38}$$

(where as usual I am using the same notation for an operator and for the matrix representing that operator). In this example, the characteristic equation gives

$$\omega(\omega - 2)^2 = 0 \tag{2.39}$$

so the eigenvectors of Ω have eigenvalues $\omega = 0, 2, 2$, and so $\omega = 2$ is degenerate. If $\omega = 0$, we find the eigenstate

$$|\omega = 0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \tag{2.40}$$

(where the factor of $1/\sqrt{2}$ is there to normalize the magnitude of the eigenvector to 1), whereas if $\omega = 2$ we find that all vectors of the form

$$\begin{pmatrix} a \\ b \\ a \end{pmatrix} \tag{2.41}$$

are eigenvectors of Ω with $\omega = 2$. As advertised, there are an infinite number of them. Thus, there is a two-dimensional subspace of our three-dimensional vector space with $\omega = 2$. We can therefore just choose two orthogonal vectors in this subspace to get orthogonal eigenvectors. Arbitrarily choosing $a = b$ gives

$$|\omega = 2\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{2.42}$$

while the other eigenvectors is chosen to be orthogonal to this

$$|\omega = 2'\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}. \tag{2.43}$$

Thus, even though we have degenerate eigenvalues, we can still choose a basis for the three-dimensional vector space consisting of orthonormal eigenvectors of Ω.

We generally refer to finding the eigenvalues and eigenvectors of a Hermitian matrix as diagonalizing the matrix, since

$$\langle \omega_j | \Omega | \omega_i \rangle = \omega_i \langle \omega_j | \omega_i \rangle = \omega_i \delta_{ij} \tag{2.44}$$

so that in the basis of eigenstates

$$|\omega_i\rangle \leftrightarrow \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \leftarrow i\text{'th place} \tag{2.45}$$

the matrix representing Ω is just the diagonal matrix consisting of its eigenvalues

$$\Omega = \begin{pmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_n \end{pmatrix} \tag{2.46}$$

You can also see this explicitly by constructing the matrix U which relates the eigenvectors $\{|\omega_i\rangle\}$ to the original basis vectors $\{|i\rangle\}$ in Eq. (2.23); for Ω defined in Eq. (2.37) we have

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}. \tag{2.47}$$

By explicitly calculating $U^\dagger \Omega U$ you can verify that Ω is diagonal in the basis of its eigenvectors.

Another useful property of Hermitian matrices is that if two Hermitian matrices *commute*, that is, if

$$\Omega\Lambda = \Lambda\Omega \qquad (2.48)$$

then there exists a basis of eigenvectors which are eigenvectors of both operators. For the case of nondegenerate eigenvalues, this is very simple to prove. By definition,

$$\Omega|\omega_i\rangle = \omega_i|\omega_i\rangle \qquad (2.49)$$

and so

$$\Lambda\Omega|\omega_i\rangle = \omega_i\Lambda|\omega_i\rangle. \qquad (2.50)$$

But since Ω and Λ commute, we can rearrange them on the left hand side to get

$$\Lambda\left(\Omega|\omega_i\rangle\right) = \omega_i\left(\Lambda|\omega_i\rangle\right). \qquad (2.51)$$

Thus, $\Lambda|\omega_i\rangle$ is an eigenvector of Ω with eigenvalue ω_i. Since the eigenvalues are nondegenerate, this means that $\Lambda|\omega_i\rangle \propto |\omega_i\rangle$, or

$$\Lambda|\omega_i\rangle = \lambda_i|\omega_i\rangle \qquad (2.52)$$

for some λ_i, and so $|\omega_i\rangle$ is also an eigenstate of Λ, with eigenvalue λ_i.

For degenerate eigenvalues, the discussion is a bit more involved, but the bottom line is that you can always choose a basis of eigenvectors common to both operators. It is more involved because, as we saw before, if you have degenerate eigenvalues then you have an entire subspace of eigenvectors; demanding that the eigenvectors also be eigenvectors of another, commuting, operator, then restricts your choice of eigenvectors in the subspace. Many books show this in full generality; here we will just illustrate with a simple example (which generalizes to more complex systems).

Consider the operator Ω defined in Eq. (2.37) and the additional operator

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad (2.53)$$

Since $[\Omega, \Lambda] = 0$, we know we can always construct a basis of eigenvectors common to both operators.

We have already found the eigenvectors of Ω; for Λ, it is easy to show that the eigenvalues of $\lambda = 0, 1, 1$ - that is, the eigenvalue 1 is doubly degenerate. The eigenvector corresponding to $\lambda = 0$ is

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad (2.54)$$

while the degenerate eigenvalue $\lambda = 2$ corresponds to any vector of the form

$$\begin{pmatrix} a \\ 0 \\ b \end{pmatrix}. \tag{2.55}$$

But now it looks like we have a problem - the states $|\omega = 2\rangle$ and $|\omega = 2'\rangle$ in Eqs. (2.42, 2.43) are therefore *not* eigenvectors of Λ! Did the theorem fail?

No - we chose the eigenstates (2.42, 2.43) arbitrarily out of the full two dimensional subspace in Eq. (2.41). Unfortunately, the eigenstates of Ω that we arbitrarily chose are not also eigenstates of Λ. So we should go back and see if we can choose eigenstates in that two dimensional subspace which are also eigenstates of Λ. After a bit of inspection, we arrive at the following states, labeled their eigenvalues ω and λ of Ω and Λ, respectively:

$$|\omega, \lambda\rangle = |2, 0\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$|\omega, \lambda\rangle = |0, 1\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \tag{2.56}$$

$$|\omega, \lambda\rangle = |2, 1\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

It is easy to see that these satisfy all the criteria: demanding that they be eigenstates of both operators picks out one particular set of eigenvectors in each degenerate subspace. Thus, even in the case of degenerate eigenvalues, we can always choose an orthonormal basis of eigenvectors common to both. The eigenvectors are then labelled by the corresponding eigenvalue for both observables. If we had larger matrices and two observables were not enough to uniquely label the eigenstates, we would have to introduce a third or more operators to uniquely define a set of basis states.

A set of observables whose eigenvectors uniquely label a complete set of basis states is known as a "complete set of commuting observables", or C.S.C.O.

3. Generalization to Infinite Dimensions

3.1 The $\{|x\rangle\}$ Basis

I mentioned earlier that the space of functions $f(x)$ defined on some interval along the real line forms a vector space - it's easy to check that the space of functions obeys all the postulates in Section 1.1. Thus, everything that we have discussed for more familiar finite-dimensional vectors spaces should carry over to the space of functions on the real line. To make things a little more familiar, let's discuss this problem as a limiting procedure.

Suppose we have a function $f(x)$ defined on the interval $0 \leq x \leq L$; the function $f(x)$ could, for example, describe the displacement of a string of length L. Rather than give the
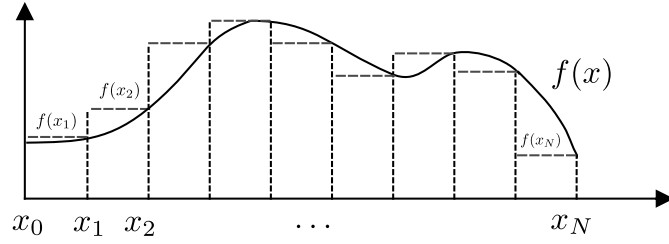
Figure 3: Approximating a function $f(x)$ by its values at discrete points x_n.

value of $f(x)$ at every point along the line, we could approximate the function by dividing the interval $0 \leq x \leq L$ into N intervals $x_1, \ldots x_N$, where $x_m = mL/N$, as shown in the Figure, and then approximate the function by

$$f_N(x) \equiv \begin{cases} f(x_1), & x_0 \leq x < x_1 \\ f(x_2), & x_1 \leq x < x_2, \\ \vdots \\ f(x_N), & x_{N-1} \leq x \leq x_N. \end{cases} \tag{3.1}$$

We can then think of the N numbers $f(x_1), \ldots f(x_N)$ as the components of a vector $|f_N\rangle$

$$|f_N\rangle = \sum_{i=0}^{N} f_i |x_i\rangle. \tag{3.2}$$

Once again, it's easy to see that the set of vectors $|f_N\rangle$ form an N-dimensional vector space. The i'th basis vector $|x_i\rangle$ is then a function which is unity for $x_{i-1} < x \leq x_i$ and zero everywhere else. We have a natural definition of the inner product of these vectors

$$\langle x_i | x_j \rangle = \delta_{ij} \tag{3.3}$$

which makes them an orthonormal basis. In addition, you can check that they obey the completeness relation

$$\sum_i |x_i\rangle\langle x_i| = \mathbb{1}. \tag{3.4}$$

The inner product of two vectors $|f_N\rangle$ and $|g_N\rangle$ is then given by

$$\langle f_N | g_N \rangle = \sum_{i=0}^{N} f^*(x_i) g(x_i) \tag{3.5}$$

and in particular the inner product of $|f_N\rangle$ with the i'th basis vector just gives $f(x_i)$:

$$\langle i | f_N \rangle = f(x_i). \tag{3.6}$$

We could also associate with each function a column vector

$$|f_N\rangle \leftrightarrow \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_N) \end{pmatrix} \tag{3.7}$$

so each basis vector corresponds to a column vector in which all components except one are zero:

$$|x_i\rangle \leftrightarrow \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{'th place} \tag{3.8}$$

and we can write the inner product of two vectors as matrix multiplication, exactly as in Eq. (1.25).

Now let's represent the function $f(x)$ exactly, by taking the limit $N \to \infty$. This gives us an infinite dimensional vector space. Unfortunately, with our current definition of the inner product, the inner product of two functions will also go to infinity. So let's just multiply our definition of the inner product by an N-dependent factor to ensure a smooth limit[9]

$$\langle f_N | g_N \rangle = \Delta \sum_{i=0}^{N} f^*(x_i) g(x_i) \tag{3.9}$$

where $\Delta \equiv L/(N+1)$. Now the $N \to \infty$ limit is smooth and just gives us the usual definition of the integral

$$\langle f | g \rangle = \int_0^L f^*(x) g(x) \, dx \tag{3.10}$$

More generally, if we define f and g on some interval $a \leq x \leq b$, the inner product will be

$$\langle f | g \rangle = \int_a^b f^*(x) g(x) \, dx. \tag{3.11}$$

We can then take the continuum limit of Eq. (3.2) to write

$$|f\rangle = \int_a^b f(x) |x\rangle dx \tag{3.12}$$

where once again the function $f(x)$ can now be interpreted as the (infinite number of) components of the vector $|f\rangle$ expressed in terms of the basis vectors $\{|x\rangle\}$. If the function is defined along the entire real line, we take the limits $a, b \to \pm\infty$ to obtain

$$|f\rangle = \int_{-\infty}^{\infty} f(x) |x\rangle dx \tag{3.13}$$

A basis vector $|x_0\rangle$ in this space then corresponds to a function which is zero everywhere except at $x = x_0$. The orthogonality requirement is then

$$\langle x | x' \rangle = 0, \ x \neq x'. \tag{3.14}$$

---

[9]This also obeys all the axioms for an inner product, so is a perfectly good definition of an inner product.

When $x = x'$, we have to be a bit careful. Let's start with the completeness relation

$$\int_a^b |x'\rangle\langle x'|dx' = \mathbb{1}. \tag{3.15}$$

Now sandwich this relation between the bra $\langle x|$ and the ket $|f\rangle$ to get

$$\int_a^b \langle x|x'\rangle\langle x'|f\rangle dx' = \langle x|f\rangle. \tag{3.16}$$

But, as with the finite-dimensional case (3.6), the inner product of $|x'\rangle$ with $|f\rangle$ is $f(x')$:

$$\langle x'|f\rangle = f(x') \tag{3.17}$$

so we have

$$\int_a^b \langle x|x'\rangle f(x')dx' = f(x). \tag{3.18}$$

Now, from Eq. (3.14), the inner product $\langle x'|x\rangle$ vanishes unless $x = x'$; thus, it is only nonzero in an infinitesimal region about $x = x'$. In this region we have $f(x') = f(x)$, so we can pull it out of the integral to give

$$f(x) \int_{x-\epsilon}^{x+\epsilon} \langle x|x'\rangle dx' = f(x) \tag{3.19}$$

(where the limit $\epsilon \to 0^+$ is implicit), or

$$\int_{x-\epsilon}^{x+\epsilon} \langle x|x'\rangle = 1. \tag{3.20}$$

Thus, the inner product $\langle x|x'\rangle$ is a very strange function - it is zero except for an infinitesimally small region about $x = x'$, and its integral is unity over any region containing $x = x'$, which means it is infinite (in a very precisely defined way) at $x = x'$. This function, which you are probably already familiar with, is called the *Dirac Delta function*:

$$\langle x|x'\rangle = \delta(x - x'). \tag{3.21}$$

The Dirac δ-function is clearly a very singular object (in fact, it's not strictly speaking a function, according to mathematicians, but a "distribution"), which obeys

$$\delta(x) = 0, \text{ if } x \neq 0$$
$$\int_a^b \delta(x)dx = 1, \text{ if } a < 0 < b. \tag{3.22}$$

That is, it is infinitely spikey at $x = 0$ such that it has a finite (unit) area. You can think of it as the limit of a Gaussian of width Δ and height $1/\Delta$ in the limit $\Delta \to 0$:

$$\delta(x) \simeq \lim_{\Delta \to 0} \frac{1}{\sqrt{\pi}\Delta} e^{-\frac{x^2}{\Delta^2}}. \tag{3.23}$$

Even though $\delta(x)$ is singular, you can manipulate it like any other function. It has some useful properties

$$\delta(x) = \delta(-x), \ (\delta(x) \text{ is even})$$

$$\delta(ax) = \frac{1}{|a|}\delta(x), \text{ or more generally,}$$

$$\delta\left(f(x)\right) = \sum_i \frac{\delta(x - x_i)}{\left|\frac{df}{dx}\right|_{x=x_i}}. \tag{3.24}$$

where the sum in the last line is over the zeroes x_i of f: $f(x_i) = 0$. You can also take derivatives of δ-functions:

$$\delta'(x - x') = \frac{d}{dx}\delta(x - x') = -\frac{d}{dx'}\delta(x - x'). \tag{3.25}$$

Derivatives of δ-functions sound even scarier than δ functions, but again since they only really make sense inside integrals, let's look at

$$\int_a^b \delta'(x - x')f(x')dx' = -\int_a^b \left[\frac{d}{dx'}\delta(x - x')\right]f(x')dx' \tag{3.26}$$

where $a < x < b$. Integrating this expression by parts, we get

$$\int_a^b \delta'(x - x')f(x')dx' = -\left.\delta(x - x')f(x')\right|_a^b + \int_a^b \delta(x - x')\frac{df(x')}{dx'}dx'$$

$$= 0 + \frac{df(x)}{dx} \tag{3.27}$$

where in the last line we have used the fact that $\delta(x)$ vanishes on the boundaries $x = a$ and $x = b$. Thus, we can write

$$\delta'(x - x') = \delta(x - x')\frac{d}{dx'} \tag{3.28}$$

where it is implicit that this equation only makes sense when integrated with respect to some function $f(x')$, which the derivative acts on.

Another useful way to represent a delta function is via its Fourier transform (see Appendix A for a brief review). Give a function $f(x)$, we define its Fourier transform by

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ikx}f(x)dx. \tag{3.29}$$

Given $\tilde{f}(k)$ we can recover the original function $f(x)$ via the inverse Fourier transform:

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{ikx}\tilde{f}(k)dk. \tag{3.30}$$

Physically, Eq. (3.30) is telling you that you can write any function $f(x)$ is a sum (or rather, integral) of an infinite number of complex exponentials (which we will also refer

to as *plane waves*) – $\tilde{f}(k)$ tells you how much of each plane wave you have. The Fourier transform of a delta function is particularly simple: if $f(x) = \delta(x - x_0)$, then

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \delta(x - x_0) dx = \frac{1}{\sqrt{2\pi}} e^{-ikx_0}. \tag{3.31}$$

Thus, the Fourier transform of a delta function is a plane wave. Taking the inverse Fourier transform, we then find

$$\delta(x - x_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x - x_0)} dk. \tag{3.32}$$

This is an interesting equation – it tells us that a delta function can be constructed by a (continuous) sum of an infinite number of plane waves. If $x \neq x_0$, this equation tells us that adding an infinite number of oscillators together gives precisely *zero* (!) - essentially, all the oscillators destructively interfere[10]. The only value for x for which the oscillators don't interfere is if $x = x_0$, in which case the exponential just becomes unity, and we get

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk \tag{3.33}$$

which is clearly divergent, and gives the ill-defined quantity $\delta(0)$. We will make extensive use of the representation (3.32) of the delta function as we proceed.

Going back now to our original expression Eq. (3.13), we can represent the quantum states of a particle on a line by the infinite dimensional state vector

$$|\psi\rangle = \int_{-\infty}^{\infty} \psi(x') |x'\rangle dx'. \tag{3.34}$$

Taking the inner product with a fixed basis vector $\langle x|$ gives[11]

$$\langle x|\psi\rangle = \int_{-\infty}^{\infty} \psi(x') \langle x|x'\rangle dx' = \psi(x). \tag{3.35}$$

Thus, the familiar wave function of quantum mechanics can be thought of as the infinite number of components of the state vector $|\psi\rangle$ in the $|x\rangle$ basis.

In particular, the wave functions for the basis vectors $\{|x\rangle\}$ are delta functions: if $|\psi\rangle = |x'\rangle$, then the wave function of the basis vector is

$$\psi(x) = \langle x|x'\rangle = \delta(x - x'). \tag{3.36}$$

As a technical aside, we note that the basis states $|x\rangle$ are *not* in fact elements of the Hilbert space of functions we are studying, because they are not square integrable. Indeed, the norm of a basis vector $|x\rangle$ is

$$\langle x|x\rangle = \int_{-\infty}^{\infty} |\delta(x)|^2 dx = \delta(0) \tag{3.37}$$

---

[10] This can be rigorously shown via an appropriate limiting procedure.
[11] This equation is the infinite dimensional analogue of $v_i = \langle i|V\rangle$.

which is undefined. Cohen-Tannoudji, et. al., make a big deal of this in Chapter 2 when they discuss using basis vectors that don't belong to the Hilbert space itself. Shankar's text deals with this issue by defining the "physical Hilbert space" consisting of both square integrable functions $f(x)$ and basis vectors normalized to delta functions, which is contrasted with the "mathematical" Hilbert space. We won't worry much about this issue - the end result is that we can still use these basis vectors to expand our functions.

3.2 The $\{|k\rangle\}$ Basis

So, we have succeeded in representing a function as a vector in an infinite dimensional vector space in a form analogous to Eq. (1.9) for finite dimensional spaces:

$$|f\rangle = \int_{-\infty}^{\infty} f(x)|x\rangle \tag{3.38}$$

Thus, we can think of the values of $f(x)$ at each point x as forming the infinite components of the vector in the basis of the $|x\rangle$'s. But now we have all the tools of linear algebra at our disposal! In particular, we can do a change of basis to some *other* set of basis vectors, and write $|f\rangle$ in terms of these new basis vectors.

A particular useful basis is the basis of complex exponentials – that is, of functions e^{ikx}. As mentioned before, we often refer to this as the plane wave basis. Consider a set of basis states $\{|k\rangle\}$, which satisfy

$$\langle x|k\rangle = \frac{1}{\sqrt{2\pi}}e^{ikx} \tag{3.39}$$

(again, remember that $\langle x|\psi\rangle$ gives the wavefunction $\psi(x)$; thus, $\langle x|k\rangle$ gives the wavefunctions of the basis states $\{|k\rangle\}$). Thus, these basis functions are complex exponentials - we will usually refer to them as *plane wave* states.[12] We can easily check that these states are normalized to δ-functions, just like the $\{|x\rangle\}$ basis:

$$\begin{aligned} \langle k'|k\rangle &= \langle k'|\mathbb{1}|k\rangle = \int_{-\infty}^{\infty} \langle k'|x\rangle\langle x|k\rangle dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ik'x}e^{ikx}dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(k'-k)x}dx = \delta(k'-k). \end{aligned} \tag{3.40}$$

They also obey the completeness relation

$$\int_{-\infty}^{\infty} |k\rangle\langle k|dk = \mathbb{1} \tag{3.41}$$

and so we can write

$$|\psi\rangle = \int_{-\infty}^{\infty} |k\rangle\langle k|\psi\rangle dk = \int_{-\infty}^{\infty} \tilde{\psi}(k)|k\rangle dk \tag{3.42}$$

where $\tilde{\psi}(k)$ is the wavefunction of the particle in the $\{|k\rangle\}$ basis (or equivalently the infinite set of components of $|\psi\rangle$ in that basis). As the notation suggests, $\tilde{\psi}(k)$ is just the Fourier transform of $\psi(x)$, which we can prove by taking the inner product $\langle x|\psi\rangle$:

$$\langle x|\psi\rangle = \psi(x) = \int_{-\infty}^{\infty} \tilde{\psi}(k)\langle x|k\rangle dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\psi}(k)e^{ikx}. \tag{3.43}$$

---

[12]We will see in lecture that these states are eigenstates of the momentum operator, so we can also refer to them as momentum eigenstates.

3.3 Linear Operators in Infinite Dimensions

As in finite dimensions, operators in infinite dimensional Hilbert spaces take states to other states

$$\Lambda|\psi\rangle = |\psi'\rangle. \tag{3.44}$$

An operator which will be important for us is the familiar differential operator, which takes a function $f(x)$ to its derivative, df/dx. Let's call this operator D. In terms of kets,

$$D|f\rangle = |df/dx\rangle \tag{3.45}$$

and so

$$\langle x|D|f\rangle = \frac{\partial f}{\partial x}. \tag{3.46}$$

We can immediately use this to work out the matrix elements of D in the x basis: or, in terms of matrix elements,

$$\langle x|D|f\rangle = \int \langle x|D|x'\rangle\langle x'|f\rangle = \frac{df(x)}{dx} \tag{3.47}$$

and so

$$\langle x|D|x'\rangle \equiv D_{xx'} = \delta'(x-x') = \delta(x-x')\frac{d}{dx'} \tag{3.48}$$

from Eq. (3.27). Now let's find the eigenvalues and eigenvectors of D. But first, let's note that D is not Hermitian: taking its adjoint, we get

$$(D_{xx'})^\dagger = D^*{}_{x'x} = \delta'(x'-x) = -\delta'(x-x') \tag{3.49}$$

so it picks up a minus sign. So we can easily get a Hermitian operator by multiplying it by $-i$:

$$K \equiv -iD \rightarrow (K_{xx'})^\dagger = K^*_{x'x} = i\delta'(x'-x) = -i\delta'(x-x') = K_{xx'}. \tag{3.50}$$

You might think that the eigenvalue problem for K is nasty since it sounds like you are diagonalizing an infinite dimensional matrix, but remember that in practice it just corresponds to finding functions which, when operated on by the derivative operator, give a multiple of the function. We already know what kind of function does that - exponentials. Indeed, we've already introduced complex exponentials in Eq. (3.39), and it is easy to verify that they indeed are eigenvectors of K:

$$\begin{aligned}
\langle x|K|k\rangle &= \int \langle x|K|x'\rangle\langle x'|k\rangle dx' \\
&= -\int i\delta'(x-x') \times \frac{1}{\sqrt{2\pi}}e^{ikx'} \\
&= \int i\delta(x-x')(-ik)\frac{1}{\sqrt{2\pi}}e^{ikx'} \\
&= k\langle x|k\rangle
\end{aligned} \tag{3.51}$$

from which we see that plane wave states $|k\rangle$ are eigenstates of K with eigenvalue k. Matrix elements of K are therefore trivial in the $\{|k\rangle\}$ basis:

$$\langle k'|K|k\rangle = k\langle k'|k\rangle = k\delta(k-k') \tag{3.52}$$

Thus, the plane wave basis $\{|k\rangle\}$ consists of normalized orthogonal eigenvectors of the Hermitian operator K. Later on in class we will see that the *momentum* operator is given by $P = \hbar K$; hence as mentioned earlier, plane-wave states are also momentum eigenstates.

What about the $\{|x\rangle\}$ basis? What operator are these states eigenvectors of? This is even easier. Define the position operator X by its action on the $|x\rangle$ basis vectors:

$$X|x\rangle = x|x\rangle \tag{3.53}$$

The matrix elements of X are then

$$\langle x'|X|x\rangle = x\langle x'|x\rangle = x\delta(x-x'). \tag{3.54}$$

We know how K acts of functions $f(x)$. How about X? If $X|f\rangle = |f'\rangle$, then

$$\langle x|f'\rangle = \langle x|X|f\rangle = \int \langle x|X|x'\rangle\langle x'|f\rangle dx' = \int x\delta(x-x')f(x')= xf(x) \tag{3.55}$$

so $f'(x) = xf(x)$, or

$$X|f\rangle = |xf\rangle. \tag{3.56}$$

This just tells us the rather obvious fact that the X operator corresponds to multiplying a function $f(x)$ by x.

Finally, we can take the matrix elements of X in the $\{|k\rangle\}$ basis:

$$\begin{aligned}
\langle k|X|k'\rangle =&\frac{1}{2\pi}\int e^{-ikx}x\, e^{ik'x}dx \\
=&\frac{i}{2\pi}\frac{d}{dk}e^{i(k'-k)x}dx \\
=&i\delta'(k-k')
\end{aligned} \tag{3.57}$$

and so

$$X|g(k)\rangle = i|dg/dk\rangle. \tag{3.58}$$

Combining this with our other matrix elements, we see that

1. In the $\{|x\rangle\}$ basis, X acts like x and K acts like $-id/dx$, and

2. In the $\{|k\rangle\}$ basis, K acts like k and X acts like id/dk.

Finally, X and K clearly do not have common eigenstates, so its not surprising that they don't commute. Lets see this explicitly, by seeing how X and K act on an arbitrary ket $|f\rangle$ in the $\{|x\rangle\}$ basis:

$$\begin{aligned}
\langle x|X|f\rangle &= xf(x) \\
\langle x|K|f\rangle &= -i\frac{df(x)}{dx}
\end{aligned} \tag{3.59}$$

and so

$$\langle x|XK|f\rangle = -ix\frac{df(x)}{dx}. \tag{3.60}$$

Similarly,

$$\langle x|KX|f\rangle = -i\frac{df}{dx}\left(xf(x)\right) = -ix\frac{df(x)}{dx} - if(x) \tag{3.61}$$

and hence

$$\langle x|[X,K]|f\rangle = i\langle x|f\rangle. \tag{3.62}$$

Since this is true for any $|f\rangle$ and $|x\rangle$, we must have

$$[X,K] = i\mathbb{1}. \tag{3.63}$$

This is a famous equation in quantum mechanics: later on, when we introduce the momentum operator $P = \hbar K$, it tells us that the position and momentum operators don't commute:

$$[X,P] = i\hbar \tag{3.64}$$

and thus, by our general discussion of noncommuting operators, these operators cannot be diagonalized simultaneously: eigenvectors of X are not eigenvectors of K (or P), as we have seen explicitly in this section.

4. Tensor Product Spaces

4.1 Two qubits

Let's return now to finite-dimensional vector spaces. The first example we studied in this course, the qubit, corresponded to a particle in a two dimensional state space spanned by the basis vectors $|\chi_\uparrow\rangle$ and $|\chi_\downarrow\rangle$:

$$|\psi\rangle = c_\uparrow|\chi_\uparrow\rangle + c_\downarrow|\chi_\downarrow\rangle \tag{4.1}$$

which, in the $\{|\chi_\uparrow\rangle, |\chi_\downarrow\rangle\}$ basis is denoted

$$|\psi\rangle \leftrightarrow \begin{pmatrix} c_\uparrow \\ c_\downarrow \end{pmatrix}. \tag{4.2}$$

Now suppose we had a second qubit[13]. Let's denote the basis vectors which span the space of the second qubit by $|\xi_\uparrow\rangle$ and $|\xi_\downarrow\rangle$. Clearly, the Hilbert space of the second qubit is also two-dimensional.

Next, suppose we want to describe the system of *both* qubits together. What is the Hilbert space which describes the combined system? Let us denote the state where the first qubit is in the state $|\chi_\uparrow\rangle$ and the second in the state $|\xi_\uparrow\rangle$ by $|\chi_\uparrow\rangle \otimes |\xi_\uparrow\rangle$. Then the Hilbert space is clearly spanned by four basis vectors,

$$|\chi_\uparrow\rangle \otimes |\xi_\uparrow\rangle, \ |\chi_\uparrow\rangle \otimes |\xi_\downarrow\rangle, \ |\chi_\downarrow\rangle \otimes |\xi_\uparrow\rangle, \ |\chi_\downarrow\rangle \otimes |\xi_\downarrow\rangle. \tag{4.3}$$

---

[13]We will assume the two qubits are distinguishable - for example, one of them could be the spin of an electron and the second the spin of a proton. For indistinguishable particles, quantum mechanics imposes some additional constraints on the allowed states which we won't consider now.

For simplicity these states are often written $|\chi_\uparrow, \xi_\uparrow\rangle \equiv |\chi_\uparrow\rangle \otimes |\xi_\uparrow\rangle$, $|\chi_\uparrow, \xi_\downarrow\rangle \equiv |\chi_\uparrow\rangle \otimes |\xi_\downarrow\rangle$, etc.

The combined system is therefore described by a four-dimensional Hilbert space spanned by the basis vectors (4.3). An arbitrary state can be represented by the ket

$$
\begin{aligned}
|\psi\rangle &= c_{\uparrow\uparrow} \left[|\chi_\uparrow\rangle \otimes |\xi_\uparrow\rangle \right] + c_{\uparrow\downarrow} \left[|\chi_\uparrow\rangle \otimes |\xi_\downarrow\rangle \right] + c_{\downarrow\uparrow} \left[|\chi_\downarrow\rangle \otimes |\xi_\uparrow\rangle \right] + c_{\downarrow\downarrow} \left[|\chi_\downarrow\rangle \otimes |\xi_\downarrow\rangle \right] \\
&\equiv c_{\uparrow\uparrow} |\chi_\uparrow, \xi_\uparrow\rangle + c_{\uparrow\downarrow} |\chi_\uparrow, \xi_\downarrow\rangle + c_{\downarrow\uparrow} |\chi_\downarrow, \xi_\uparrow\rangle + c_{\downarrow\downarrow} |\chi_\downarrow, \xi_\downarrow\rangle.
\end{aligned}
\tag{4.4}
$$

The Hilbert space of the combined system is called the *tensor product space* of the state spaces of the two individual qubits. In the $\{|\chi_\uparrow, \xi_\uparrow\rangle, |\chi_\uparrow, \xi_\downarrow\rangle, |\chi_\downarrow, \xi_\uparrow\rangle, |\chi_\downarrow, \xi_\downarrow\rangle\}$ basis we can represent the components of $|\psi\rangle$ by a four-dimensional column vector in the tensor product space

$$
|\psi\rangle \leftrightarrow \begin{pmatrix} c_{\uparrow\uparrow} \\ c_{\uparrow\downarrow} \\ c_{\downarrow\uparrow} \\ c_{\downarrow\downarrow} \end{pmatrix}.
\tag{4.5}
$$

Tensor product spaces tend to behave in the way you would expect. For example, the tensor product of two states is linear and distributive. This means if we have a state where the first particle is in the superposition $a_1|\chi_\uparrow\rangle + a_2|\chi_\downarrow\rangle$ and the second particle is in $|\xi_\downarrow\rangle$, the two-particle system is in the state

$$
|\psi\rangle = (a_1|\chi_\uparrow\rangle + a_2|\chi_\downarrow\rangle) \otimes |\xi_\downarrow\rangle = a_1 \left[|\chi_\uparrow\rangle \otimes |\xi_\downarrow\rangle \right] + a_2 \left[|\chi_\downarrow\rangle \otimes |\xi_\downarrow\rangle \right] = a_1|\chi_\uparrow, \xi_\downarrow\rangle + a_2|\chi_\downarrow, \xi_\downarrow\rangle.
\tag{4.6}
$$

Furthermore, since the rules of quantum mechanics don't care whether we are working with a tensor product space or not, all the usual rules hold: for example, in (4.4), $|c_{\uparrow\uparrow}|^2$ is the probability to find the system in the state $|\chi_\uparrow, \xi_\uparrow\rangle$ - that is, to find the first qubit in the state $|\chi_\uparrow\rangle$ and the second in $|\xi_\uparrow\rangle$ (assuming the state is normalized).

Example: If a system is in the state (4.4), what are the probabilities to find (a) the first qubit in the state $|\chi_\uparrow\rangle$, (b) the second qubit in the state $|\xi_\downarrow\rangle$, and (c) the first qubit in the state $|\chi_\uparrow\rangle$ and the second in the state $|\xi_\rightarrow\rangle \equiv \frac{1}{\sqrt{2}} (|\xi_\uparrow\rangle + |\xi_\downarrow\rangle)$?
Answers: (a) $|\langle\chi_\uparrow, \xi_\uparrow|\psi\rangle|^2 + |\langle\chi_\uparrow, \xi_\downarrow|\psi\rangle|^2 = |c_{\uparrow\uparrow}|^2 + |c_{\uparrow\downarrow}|^2$, (b) $|\langle\chi_\uparrow, \xi_\downarrow|\psi\rangle|^2 + |\langle\chi_\downarrow, \xi_\downarrow|\psi\rangle|^2 = |c_{\uparrow\downarrow}|^2 + |c_{\downarrow\downarrow}|^2$, (c) $|\langle\chi_\uparrow, \xi_\rightarrow|\psi\rangle|^2 = \frac{1}{2} |\langle\chi_\uparrow, \xi_\uparrow|\psi\rangle + \langle\chi_\uparrow, \xi_\downarrow|\psi\rangle|^2 = \frac{1}{2} |c_{\uparrow\uparrow} + c_{\uparrow\downarrow}|^2$.

4.2 Operators in Tensor Product Spaces

The previous discussion easily generalizes to tensor products of arbitrary vector spaces. If we have two particles where the state of the first particle is described by the vector space $\mathbb{V}$ and the state of the second is described by the vector space $\mathbb{W}$, then we denote the state of the two-particle system by the tensor product space denoted $\mathbb{V} \otimes \mathbb{W}$. If $\mathbb{V}$ is an n-dimensional vector space and $\mathbb{W}$ is an m-dimensional vector space, then $\mathbb{V} \otimes \mathbb{W}$ is an $n \times m$ dimensional vector space. If we have k particles where the i'th particle is described by the n_i-dimensional Hilbert space $\mathbb{V}_i$, we can define the $n_1 \times n_2 \times \cdots \times n_k$ dimensional tensor product space $\mathbb{V}_1 \otimes \mathbb{V}_2 \otimes \cdots \otimes \mathbb{V}_k$ in an analogous way.

Now let's consider the action of operators which act on individual vector spaces on the tensor product. Suppose we have two vector spaces $\mathbb{V}$ and $\mathbb{W}$, respectively, and a linear

operator T which acts on $\mathbb{V}$ and another operator S which acts on $\mathbb{W}$. There is then a natural way to combine T and S to construct an operator $T \otimes S$ which operators on the two-particle state space $\mathbb{V} \otimes \mathbb{W}$:

$$(T \otimes S) \left[|\psi_1\rangle \otimes |\psi_2\rangle \right] = [T|\psi_1\rangle] \otimes [S|\psi_2\rangle] \tag{4.7}$$

where $|\psi_1\rangle \in \mathbb{V}$ and $|\psi_2\rangle \in \mathbb{W}$. As you would expect, in the tensor product $T \otimes S$, T acts on the first state and S acts on the second.

If we have an operator which just acts on one of the vector spaces, then it doesn't do anything to vectors in the other vector space, which means it is just the identity on the other space. Thus, if T acts only on V, we have

$$T \left[|\psi_1\rangle \otimes |\psi_2\rangle \right] = (T \otimes \mathbb{1}) \left[|\psi_1\rangle \otimes |\psi_2\rangle \right] = [T|\psi_1\rangle] \otimes |\psi_2\rangle \tag{4.8}$$

(and similarly for an operator which acts only on W).

Example: The operators σ_x and σ_y were defined in (2.34) and (2.32). What is the 4×4 matrix which represents the operator $\sigma_x \otimes \sigma_y$ on the space of two qubits in the basis (4.3)? From (4.7), we can work out the effect of $\sigma_x \otimes \sigma_y$ on each of the four basis vectors. For example,

$$\begin{aligned} (\sigma_x \otimes \sigma_y) \, |\chi_\uparrow, \xi_\uparrow\rangle &= \sigma_x|\chi_\uparrow\rangle \otimes \sigma_y|\xi_\uparrow\rangle \\ &= |\chi_\downarrow\rangle \otimes (i|\xi_\downarrow\rangle) \\ &= i|\chi_\downarrow, \xi_\downarrow\rangle. \end{aligned} \tag{4.9}$$

Similarly, you should find

$$\begin{aligned} (\sigma_x \otimes \sigma_y) \, |\chi_\uparrow, \xi_\downarrow\rangle &= -i|\chi_\downarrow, \xi_\uparrow\rangle \\ (\sigma_x \otimes \sigma_y) \, |\chi_\downarrow, \xi_\uparrow\rangle &= i|\chi_\uparrow, \xi_\downarrow\rangle \\ (\sigma_x \otimes \sigma_y) \, |\chi_\downarrow, \xi_\downarrow\rangle &= -i|\chi_\uparrow, \xi_\uparrow\rangle \end{aligned} \tag{4.10}$$

and so in the basis (4.3), the operator is represented by the matrix

$$\sigma_x \otimes \sigma_y \leftrightarrow \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}. \tag{4.11}$$

4.3 Systems in Three Dimensions

Tensor products don't just arise when we have systems with multiple particles - they arise any time we want to add more degrees of freedom to a system, such as going from a particle moving in one dimension to a particle moving in three dimensions. As we know, if we have a particle confined to one dimension (say, the x direction), the state space is spanned by the basis vectors

$$\{|x\rangle\} \tag{4.12}$$

which each correspond to a particle localized at the position x. Let's call this vector space $\mathcal{E}_x$. As we have discussed, $\mathcal{E}_x$ is infinite dimensional, since it has an infinite number of orthonormal basis states. The $|x\rangle$ basis states are eigenstates of the x-position operator X:

$$X|x\rangle = x|x\rangle. \tag{4.13}$$

Similarly, if we had a particle constrained to lie along the y-axis, its Hilbert space would be spanned by the basis vectors

$$\{|y\rangle\} \tag{4.14}$$

which each correspond to a particle at the location y and are eigenstates of the y-position operator Y:

$$Y|y\rangle = y|y\rangle. \tag{4.15}$$

These basis vectors span the vector space $\mathcal{E}_y$. We could similarly define the Hilbert space $\mathcal{E}_z$ of a particle constrained to lie along the z-axis, with basis states that are eigenstates of the Z operator.

If we denote by $\mathcal{E}_{\vec{r}}$ the state space of a particle moving in three dimensions, its state space is spanned by the basis vectors

$$\{|\vec{r}\rangle\} = \{|x, y, z\rangle\} = \{|x\rangle \otimes |y\rangle \otimes |z\rangle\} \tag{4.16}$$

which correspond to a particle localized at the point $\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}$. Comparing this with (4.4) , we immediately see that $\mathcal{E}_{\vec{r}}$ is the tensor product

$$\mathcal{E}_{\vec{r}} = \mathcal{E}_x \otimes \mathcal{E}_y \otimes \mathcal{E}_z. \tag{4.17}$$

The basis states in (4.16) are eigenstates of the position operators X, Y, and Z[14]

$$X|\vec{r}\rangle = x|\vec{r}\rangle, \ Y|\vec{r}\rangle = y|\vec{r}\rangle, \ Z|\vec{r}\rangle = z|\vec{r}\rangle. \tag{4.18}$$

It's convenient to combine the operators X, Y and Z into a *vector operator* $\vec{R} \equiv X\hat{x} + Y\hat{y} + Z\hat{z}$ to write (4.18) as the eigenvalue equation

$$\vec{R}|\vec{r}\rangle = \vec{r}|\vec{r}\rangle. \tag{4.19}$$

Note that we are using the word "vector" in three different ways now! A particle in three dimensions is described by an infinite-dimensional *state vector* $|\psi\rangle$ living in the space $\mathcal{E}_{\vec{r}}$, while $\vec{r}$ is a three-dimensional *geometric vector*. We also have a *vector operator* $\vec{R}$ whose components are *operators* acting on $\mathcal{E}_{\vec{r}}$. Both $\vec{r}$ and $\vec{R}$ are three-dimensional, but the components of $\vec{r}$ are numbers x, y and z, while the components of $\vec{R}$ are operators X, Y and Z. Be careful not to confuse these three objects!

We can now work out the properties of $\mathcal{E}_{\vec{r}}$ as a straightforward extension of the one-dimensional case. The basis states of $\mathcal{E}_{\vec{r}}$ are normalized

$$\begin{aligned}
\langle \vec{r}'|\vec{r}\rangle &= \left[\langle x'| \otimes \langle y'| \otimes \langle z'|\right] | \left[|x\rangle \otimes |y\rangle \otimes |z\rangle\right] \\
&= \langle x'|x\rangle\langle y'|y\rangle\langle z'|z\rangle \\
&= \delta^{(3)}(\vec{r} - \vec{r}')
\end{aligned} \tag{4.20}$$

---

[14]Note that I am using the property (4.8): the X operator acts as the identity operator on $\mathcal{E}_y$ and $\mathcal{E}_z$, so $X|\vec{r}\rangle = X|x\rangle \otimes \mathbb{1}|y\rangle \otimes \mathbb{1}|z\rangle = x|\vec{r}\rangle$, and similarly for Y and Z.

where we define the three-dimensional delta function

$$\delta^{(3)}(\vec{r} - \vec{r}\,') \equiv \delta(x - x')\delta(y - y')\delta(z - z'). \tag{4.21}$$

We can define the momentum operator (another vector operator)[15]

$$\vec{P} \equiv \hbar\vec{K} = P_x\hat{x} + P_y\hat{y} + P_z\hat{z} \tag{4.22}$$

and momentum eigenstates

$$|\vec{p}\rangle = |p_x, p_y, p_z\rangle \equiv |p_x\rangle \otimes |p_y\rangle \otimes |p_z\rangle \tag{4.23}$$

which obey

$$P_x|\vec{p}\rangle = p_x|\vec{p}\rangle, \ \ P_y|\vec{p}\rangle = p_y|\vec{p}\rangle, \ \ P_z|\vec{p}\rangle = p_z|\vec{p}\rangle \tag{4.24}$$

or equivalently

$$\vec{P}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle. \tag{4.25}$$

These states are also normalized to delta functions:

$$\langle\vec{p}\,'|\vec{p}\rangle = \delta^{(3)}(\vec{p} - \vec{p}\,'). \tag{4.26}$$

The position and momentum operators obey the commutation relations

$$[X, P_x] = [Y, P_y] = [Z, P_z] = i\hbar \tag{4.27}$$

while all other position and momentum operators commute (i.e. $[X, Y] = 0$, $[P_x, P_z] = 0$, $[X, P_y] = 0$, etc.). We can write (4.27) compactly as

$$[R_i, P_j] = i\hbar\delta_{ij}. \tag{4.28}$$

We define the position- and momentum-space wave functions as

$$\psi(\vec{r}) \equiv \langle\vec{r}|\psi\rangle, \ \ \tilde{\psi}(\vec{p}) \equiv \langle\vec{p}|\psi\rangle. \tag{4.29}$$

The completeness relations for the $\vec{r}$ or $\vec{p}$ basis states are

$$\int d^3r|\vec{r}\rangle\langle\vec{r}| = \int d^3p|\vec{p}\rangle\langle\vec{p}| = 1 \tag{4.30}$$

and the inner product of any two states in three dimensions can be written as

$$\langle\psi'|\psi\rangle = \int d^3r \ \langle\psi'|\vec{r}\rangle\langle\vec{r}|\psi\rangle = \int d^3r \ \psi'^*(\vec{r})\psi(\vec{r}) \tag{4.31}$$

(note that we are now doing triple integrals) or equivalently, in the $\vec{p}$ basis,

$$\langle\psi'|\psi\rangle = \int d^3p \ \langle\psi'|\vec{p}\rangle\langle\vec{p}|\psi\rangle = \int d^3p \ \psi'^*(\vec{p})\psi(\vec{p}) \tag{4.32}$$

---

[15]Remember, the components of $\vec{P}$ are operators, while the components of $\vec{p}$ are numbers.

In one dimension, the eigenstates of momentum $|p\rangle$ are related to the $|k\rangle$ eigenstates we studied earlier by

$$|p\rangle = \hbar^{-1/2}|k\rangle \tag{4.33}$$

(you should check this by comparing the normalizations of $|k\rangle$ and $|p\rangle$) and so we can write

$$\begin{aligned}
\psi(x) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\ \tilde{\psi}(p) e^{ipx/\hbar} \\
\tilde{\psi}(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx\ \psi(x) e^{-ipx/\hbar}.
\end{aligned} \tag{4.34}$$

In three dimensions, this becomes

$$\begin{aligned}
\psi(\vec{r}) &= \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p\ \tilde{\psi}(\vec{p}) e^{i\vec{p}\cdot\vec{r}/\hbar} \\
\tilde{\psi}(\vec{p}) &= \frac{1}{(2\pi\hbar)^{3/2}} \int d^3r\ \psi(\vec{r}) e^{-i\vec{p}\cdot\vec{r}/\hbar}.
\end{aligned} \tag{4.35}$$

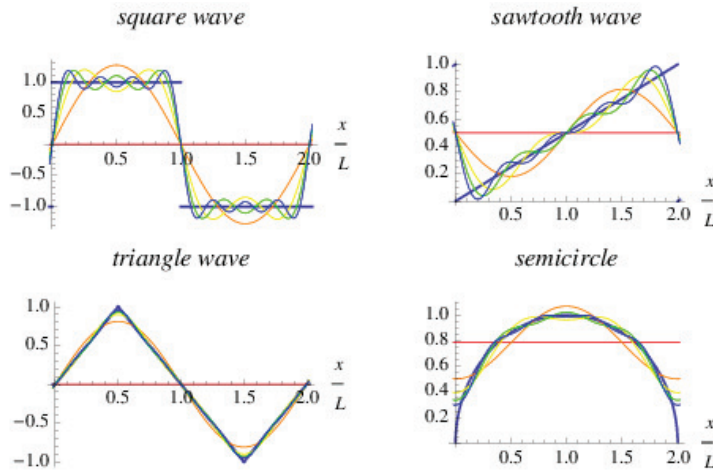## A. Fourier Series and Fourier Transforms

A **Fourier Series** lets you express any function defined on a finite interval – for simplicity we'll take from $x = -L$ to $x = +L$ – as the sum of a series of sines and cosines. For a real function $g(x)$, we have

$$g(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L} + \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L} \tag{A.1}$$

where

$$a_n = \frac{1}{L} \int_{-L}^{L} g(x) \cos \frac{n\pi x}{L} dx, \ \ b_n = \frac{1}{L} \int_{-L}^{L} g(x) \sin \frac{n\pi x}{L} dx. \tag{A.2}$$

This is illustrated in Fig. 4 for a few examples.

Figure 4: Some examples of Fourier series: adding sines and cosines with different coefficients to produce various functions (from **Wolfram Mathworld**).

For a *complex* function $f(x)$, we can do the same thing with complex exponentials $e^{ikx} = \cos kx + i \sin kx$:

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{\frac{i\pi n x}{L}}, \ \text{where} \ a_n = \frac{1}{2L} \int_{-L}^{L} f(x) e^{\frac{-i\pi n x}{L}}. \tag{A.3}$$

Thus, any complex function defined on some finite interval (or periodic over that integral) can be written as a sum of complex exponentials.

Now let's think about the expansion Eq. (A.3) in the language of linear algebra. We can think of the expansion (A.3) as the expansion of $f(x)$ on a set of *basis functions* $e^{in\pi x/L}$, and the Fourier coefficients a_n as the components of the function $f(x)$ in that basis, just like the expansion of a vector in terms of basis vectors in Eq. (1.9). To make this explicit, we can associate a ket $|f\rangle$ with the function $f(x)$ as in Eq. (3.12) to write

$$|f\rangle = \int_{-L}^{L} f(x)|x\rangle dx = \sum_{n=-\infty}^{\infty} c_n|n\rangle \tag{A.4}$$

where we have defined the basis states $|n\rangle$ with position-space wavefunctions

$$\langle x|n\rangle = \frac{1}{\sqrt{2L}}e^{in\pi x/L} \tag{A.5}$$

(the factor of $\sqrt{2L}$ is there for later convenience), and therefore $c_n = \sqrt{2L}a_n$. Eq. (A.4) then indicates that we can expand the ket $|f\rangle$ in terms of either the position-space basis $\{|x\rangle\}$ *or* the basis of complex exponentials $\{|n\rangle\}$ (note that in the $\{|x\rangle\}$ basis the basis states are continuous, while in the $\{|n\rangle\}$ basis they are discrete.)

It is easy to check that the $\{|n\rangle\}$ basis states are orthonormal:

$$\langle m|n\rangle = \frac{1}{2L}\int_{-L}^{L} e^{-im\pi x/L}e^{in\pi x/L} = \begin{cases} 0, & m \neq n \\ 1, & m = n \end{cases}. \tag{A.6}$$

(For $n \neq m$ the oscillating exponentials destructively interfere in the integral to make the total integral 0; for $n = m$ the product of the exponentials is one.)

Since the basis vectors are orthonormal we can manipulate them as usual; in particular, we find

$$a_n = \langle n|f\rangle = \langle n|\int_{-L}^{L} f(x)|x\rangle dx = \int_{-L}^{L} f(x)\langle n|x\rangle dx = \frac{1}{\sqrt{2L}}\int_{-L}^{L} f(x)e^{-in\pi x/L}dx \tag{A.7}$$

in agreement with Eq. (A.3).

It is straightforward to extend this to a function defined over the entire real line. Let's define $k = n\pi/L$, so the sum over n is equivalently a sum over k. Then in the limit $L \to \infty$, the sum over k becomes an integral $\int dk$, and the Fourier coefficients a_n are replaced by functions of k, which we'll denote $\tilde{f}(k)$. We define the **Fourier Transform** $\tilde{f}(k)$ of $f(x)$:

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ikx}f(x)dx. \tag{A.8}$$

Given $\tilde{f}(k)$, we can recover the original function $f(x)$ by the Inverse Fourier Transform

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{ikx}\tilde{f}(k)dk \tag{A.9}$$

(note that where you put the factors of $1/\sqrt{2\pi}$ is a convention.) As with the finite L case, the Inverse Fourier Transform allows us to write a function $f(x)$ as an infinite sum of complex exponentials, this time labeled by the continuous index k.

Once again, we can write this expansion in the language of linear algebra by defining basis vectors $\{|k\rangle\}$, where

$$\langle x|k\rangle = \frac{1}{\sqrt{2\pi}}e^{ikx} \tag{A.10}$$

are the plane wave states introduced in Eq. (3.39), and so $|f\rangle$ can be written in either the $\{|x\rangle\}$ basis or the $\{|k\rangle\}$ basis:

$$|f\rangle = \int_{-\infty}^{\infty} f(x)|x\rangle dx = \int_{-\infty}^{\infty} \tilde{f}(k)|k\rangle dk. \tag{A.11}$$

The wave function $\tilde{f}(k)$ in the $\{|k\rangle\}$ basis is therefore the Fourier transform of the wave-function $f(x)$ in the $\{|x\rangle\}$ basis. As with the case for finite L with discrete basis vectors, we can check the normalization of these states

$$\langle k'|k\rangle = \langle k'|\mathbb{1}|k\rangle = \int_{-\infty}^{\infty} \langle k'|x\rangle \langle x|k\rangle dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ik'x} e^{ikx} dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(k'-k)x} dx.$$
(A.12)

That last integral looks a bit suspect and possibly ill-defined, but in fact it is equal to a delta function (see Eq. (3.40)):

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(k'-k)x} dx = \delta(k'-k)$$
(A.13)

We can see this as follows: if $f(x) = \delta(x - x_0)$, then

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \delta(x - x_0) dx = \frac{1}{\sqrt{2\pi}} e^{-ikx_0}.$$
(A.14)

So the Fourier transform of a δ-function is a complex exponential. Similarly, taking the inverse Fourier transform, Eq. (A.9) of Eq. (A.14) we obtain

$$\delta(x - x_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x_0)} dk$$
(A.15)

from which Eq. (A.13) follows after relabelling the variables. (Intuitively, if $x \neq x_0$ in Eq. (A.15), you are adding together an infinite number of oscillators which destructively interfere, giving 0 for the integral. If $x = x_0$, you get $\frac{1}{2\pi} \int_{-\infty}^{\infty} dk$, which is infinite.)

If integrals like Eq. (A.15) bother you, remember that a delta function is only really defined when integrated against some sufficiently well-behaved function $f(x)$. So let's show that this equation is true when integrated against some smooth function $f(x)$ which goes to zero sufficiently rapidly at $x \to \pm\infty$. First of all, write Eq. (A.15) as the limit

$$\delta(x) = \frac{1}{2\pi} \lim_{K \to \infty} \int_{-K}^{K} e^{ikx} dx = \frac{1}{2\pi} \lim_{K \to \infty} \frac{2 \sin Kx}{x}.$$
(A.16)

Of course this limit doesn't formally exist, but if we integrate this expression with respect to $f(x)$ we get the well-defined result

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x) \delta(x) dx &= \frac{1}{\pi} \lim_{K \to \infty} \int_{-\infty}^{\infty} \frac{\sin Kx}{x} f(x) dx \\
&= \frac{1}{\pi} \lim_{K \to \infty} \int_{-\infty}^{\infty} \frac{\sin y}{y} f(y/K) dy \\
&= \frac{1}{\pi} f(0) \int_{-\infty}^{\infty} \frac{\sin y}{y} dy = f(0)
\end{aligned}$$
(A.17)

as required (where we changed variables to $y = Kx$ in the second line).