Functions as vectors

A number of times in the early sections of the course we used functions as examples of vectors. If we confine ourselves to, say, polynomials of a given order, we have a finite dimensional space. But clearly without that restriction, the order is infinite, and that introduces new issues.

In the first two sections, we will consider functions as vectors. Then in the subsequent sections we will find a way of mapping abstract vectors on to functions.

Shankar covers this material in a slightly different order. Other textbooks cover the material but within the context of quantum mechanics from the start; see eg Griffiths Chapter 3.

2.1 Inner product for functions

Shankar p 59

So far we have not defined an inner product in a function space. The definition that we will find useful is as follows. Given two complex functions of the real variable $x \in \mathbb{R}$, $f(x)$ and $g(x)$, both vectors in the space and so also written $|f\rangle$ and $|g\rangle$, then

$$\langle f|g\rangle = \int_{-\infty}^{\infty} f^*(x)g(x)\,\mathrm{d}x \qquad |f|^2 = \langle f|f\rangle = \int_{-\infty}^{\infty} f^*(x)f(x)\,\mathrm{d}x$$

It is easily seen that this satisfies the rules for an inner product; in particular as $f^*(x)f(x) \geq 0$, if $\langle f|f\rangle = 0$ then $f(x) = 0$ for all x—the zero function.

Take careful note that while f and g are functions of x, $\langle f|g\rangle$ is just a complex number, NOT a function of x.

However this inner product is not defined for all functions, only those that are **square integrable**, that is $\langle f|f\rangle$ is finite. So the space of square-integrable functions of $x \in \mathbb{R}$ is an inner-product or Hilbert space. (We note that the Schwartz inequality ensures $\langle f|g\rangle$ will be finite if f and g are square integrable, and the triangle inequality ensures a linear combination ("vector sum") of square-integrable functions is also square integrable $^{\mathbb{P}}$.)

With an eye to the application to quantum mechanics, and with due disregard for mathematical rigour, we will confine ourselves to the subspace of "well-behaved" continuous functions, for which square-integrability also implies that f vanishes as $x \to \pm\infty$. In most cases we will require $f'(x)$ and $xf(x)$ to be in the space as well; this restriction will be assumed in what follows. One exception will be if the functions are required to vanish outside a finite range of x.

Given an inner product we can find sets of functions which are orthogonal; an example is

$\{\phi_0(x) = N_0\mathrm{e}^{-x^2/2},\ \phi_1(x) = N_1 2x\mathrm{e}^{-x^2/2},\ \phi_2(x) = N_2(4x^2-2)\mathrm{e}^{-x^2/2},\ \phi_3(x) = N_3(8x^3-12x)\mathrm{e}^{-x^2/2}\}$.

(The numbers N_n are conventionally chosen to normalise the functions and make the set orthonormal.) Any finite set of course cannot be a basis, but an infinite set can; an example is the set $\{\phi_n(x) = N_n H_n(x) \mathrm{e}^{-x^2/2}\}$ where $H_n(x)$ is the nth Hermite polynomial, the first four of which $(n = 0, 1, 2, 3)$ give the previously listed set.

We will call the nth normalised member of an orthonormal basis $\phi_n(x)$ or $|n\rangle$, where by convention and depending on the basis $n = 0, 1, 2 \ldots$ or $n = 1, 2 \ldots$. So now $|0\rangle$ may represent a basis vector, NOT the zero vector, which will be written 0.

Since this set is a basis, any $f(x)$ in the space can be written $f(x) = \sum_n f_n |n\rangle$ where the infinite list of complex components $\{f_0, f_1, f_2, \ldots\}$ is the infinite-length column vector which represents $|f\rangle$ in this basis. As expected the following results hold $^{\mathbb{P}}$:

$$f_n = \langle n|f\rangle = \int_{-\infty}^{\infty} \phi_n^*(x) f(x)\, \mathrm{d}x; \qquad \langle f|g\rangle = \sum_n f_n^* g_n; \qquad \langle f|f\rangle = \sum_n |f_n|^2 < \infty.$$

2.2 Operators in function spaces

Shankar pp 63-64

Somewhat confusingly, the simplest kind of operator in function space is multiplication by another function! In particular multiplication by x will give us another function in the space. The new function $xf(x)$ is written in abstract notation as $\hat{X}|f\rangle$.

Another operator is differentiation: $\mathrm{d}f/\mathrm{d}x$ is another function in the space. In abstract notation, we write $\hat{D}|f\rangle$.

$\hat{X}$ is obviously Hermitian, since $\langle f|\hat{X}|g\rangle$ can be written $\int f^* \times (xg)\, \mathrm{d}x$, but that is equivalent to $(\int g^* \times (xf)\, \mathrm{d}x)^*$ which is $\langle g|\hat{X}|f\rangle^*$.

What about $\hat{D}$? Consider

$$\langle f|\hat{D}|g\rangle = \int_{-\infty}^{\infty} f^*(x) \frac{\mathrm{d}g}{\mathrm{d}x}\, \mathrm{d}x = \Big[f^* g\Big]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\mathrm{d}f^*}{\mathrm{d}x} g(x)\, \mathrm{d}x = -\langle g|\hat{D}|f\rangle^*$$

where we have integrated by parts and used the fact that f and g vanish at $\pm\infty$. So $\hat{D}$ is antihermitian, but $\hat{K} \equiv -i\hat{D}$ is Hermitian $^{\mathbb{P}}$. (Looking ahead, when we use these ideas in quantum mechanics we will be using $\hat{P} \equiv -i\hbar\hat{D}$ instead, but the constant is irrelevant just now.)

By integrating by parts twice we can show that $\hat{D}^2$ is Hermitian. So is $\hat{K}^2$.

From the fact that the Hermite polynomials are the solutions (with integer $n \geq 0$) of the equation

$$\frac{\mathrm{d}^2 H_n}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}H_n}{\mathrm{d}x} = -2nH_n$$

we can show that the set $\{H_n(x)\mathrm{e}^{-x^2/2}\}$ are eigenfunctions of the Hermitian operator $\hat{K}^2 + \hat{X}^2$ with eigenvalues $2n+1$ $^{\mathbb{P}}$. As expected, the eigenvalues of the Hermitian operator are real and the eigenvectors (eigenfunctions) are orthogonal. In this basis, $\hat{K}^2 + \hat{X}^2$ is represented by an infinite-dimensional diagonal square matrix with matrix elements $\langle m|\hat{K}^2 + \hat{X}^2|n\rangle = (2n+1)\delta_{mn}$.

The operators $\hat{X}$ and $\hat{D}$ do not commute: $[\hat{X}, \hat{D}] \neq 0$. If we consider an arbitrary function

$f(x)$, then

$$[\hat{X},\hat{D}]|f\rangle \equiv \hat{X}\hat{D}|f\rangle - \hat{D}\hat{X}|f\rangle \longrightarrow x\frac{\mathrm{d}f}{\mathrm{d}x} - \frac{\mathrm{d}(xf)}{\mathrm{d}x} = -f(x) \Rightarrow [\hat{X},\hat{D}] = -1$$

Equivalently, $[\hat{X},\hat{K}] = i$.

If $Q(x)$ is a polynomial in x, and $\mathrm{d}Q/\mathrm{d}x = R(x)$, we can also write down the operators $\hat{Q} = Q(\hat{X})$ and $\hat{R} = R(\hat{X})$. Then[ℙ] $[\hat{Q},\hat{X}] = 0$ and $[\hat{Q},\hat{K}] = i\hat{R}$.

2.3 Eigenstates of $\hat{X}$ and the x-representation

Shankar pp 57-70

Let us define eigenkets of $\hat{X}$, denoted $|x_0\rangle$, such that $\hat{X}|x_0\rangle = x_0|x_0\rangle$, where x_0 is a real number. Clearly x_0 can take any value at all, so there are uncountably many such kets, including $|2.5\rangle$, $|-53.34\rangle$, $|\sqrt{2}\rangle$, $|\pi\rangle$ .... Often we don't want to specify the value but keep it general, giving $\hat{X}|x\rangle = x|x\rangle$ for any x. Different eigenkets are orthogonal: $\langle x|x'\rangle = 0$ if $x \neq x'$. The set $\{|x\rangle\}$ is called the x-basis. The completeness relation for the identity now involves a sum over all these states, but a sum over a continuous variable is an integral, so we have

$$\int_{-\infty}^{\infty} |x\rangle\langle x|\,\mathrm{d}x = \hat{I}$$

We will often use x' or even x'' as the variable of integration.

Now consider $\langle x|f\rangle$. This is the x component of an abstract vector $|f\rangle$, which is a complex number that varies with x, i.e. a function of x which we can call $f(x)$. So if we haven't already specified the type of object that $|f\rangle$ is, the x-basis gives us a way of associating a function $f(x)$ with it. In this way of looking at things, $f(x)$ is the representation of $|f\rangle$ in the x-basis:

$$|f\rangle \underset{x}{\longrightarrow} f(x).$$

It follows (using the expresson above for the identity operator) that

$$|f\rangle = \int_{-\infty}^{\infty} |x\rangle\langle x|f\rangle\,\mathrm{d}x = \int_{-\infty}^{\infty} f(x)|x\rangle\,\mathrm{d}x; \qquad \langle f|g\rangle = \int_{-\infty}^{\infty} \langle f|x\rangle\langle x|g\rangle\,\mathrm{d}x \int_{-\infty}^{\infty} f^*(x)g(x)\,\mathrm{d}x;$$

with the latter equation giving the expected definition of the inner product for functions.

But what is the function associated with the ket $|x_0\rangle$, $\langle x|x_0\rangle$? We already know that it is a rather strange object: somehow it only"knows about" the specific point $x = x_0$. Consider the following:

$$f(x) = \langle x|f\rangle = \langle x|\left(\int_{-\infty}^{\infty} |x'\rangle\langle x'|\,\mathrm{d}x'\right)|f\rangle = \int_{-\infty}^{\infty} \langle x|x'\rangle f(x')\,\mathrm{d}x'$$

We should recognise this type of expression: for this to work, we must have $\langle x|x'\rangle = \delta(x-x')$, the Dirac delta function. The delta function is real and symmetric, $\delta(x-x') = \delta(x'-x)$, so as required $\langle x'|x\rangle = \langle x|x'\rangle^*$. This fixes the normalisation of the kets $|x\rangle$, and it is different from $\langle n|n\rangle = 1$, which is appropriate for a countable (discrete) basis.

The matrix elements of any operator $\hat{A}$ in the x-basis are $\langle x|\hat{A}|x'\rangle$. This is obviously a function of two variables. However many operators vanish unless $x = x'$; these are called **local**. An example is $\hat{X}$ itself: $\langle x|\hat{X}|x'\rangle = x'\delta(x-x')$ (or equivalently $x\delta(x-x')$).

Finally let us consider $\hat{D}$. We want the representation of $\hat{D}|f\rangle$ to be $|\mathrm{d}f/\mathrm{d}x\rangle$, i.e.

$$\langle x|\hat{D}|f\rangle = \frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\langle x|f\rangle.$$

Then

$$\langle x|\hat{D}|x'\rangle = \frac{\mathrm{d}}{\mathrm{d}x}\langle x|x'\rangle = \frac{\mathrm{d}}{\mathrm{d}x}\delta(x-x').$$

Note that, as expected since $\hat{D}$ is antihermitian, $(\mathrm{d}/\mathrm{d}x)\delta(x-x') = -(\mathrm{d}/\mathrm{d}x')\delta(x-x')$.

If the delta function is weird, its derivative is even weirder. But remember, both only have meaning within an integral (technically speaking, they are **distributions** rather than functions). So

$$\begin{aligned}\langle x|\hat{D}|f\rangle &= \int_{-\infty}^{\infty}\langle x|\hat{D}|x'\rangle\langle x'|f\rangle\,\mathrm{d}x' = \int_{-\infty}^{\infty}\frac{\mathrm{d}\delta(x-x')}{\mathrm{d}x}f(x')\,\mathrm{d}x' \\ &= -\int_{-\infty}^{\infty}\frac{\mathrm{d}\delta(x-x')}{\mathrm{d}x'}f(x')\,\mathrm{d}x' = \int_{-\infty}^{\infty}\delta(x-x')\frac{\mathrm{d}f}{\mathrm{d}x'}\,\mathrm{d}x' = \frac{\mathrm{d}f}{\mathrm{d}x}\end{aligned}$$

as expected.

Note $\hat{D}$ is also local.

For local operators, it is very common to drop the delta function and just write, say, $\hat{X} \underset{x}{\longrightarrow} x$, $\hat{D} \underset{x}{\longrightarrow} \mathrm{d}/\mathrm{d}x$, $\hat{K} \underset{x}{\longrightarrow} -i\mathrm{d}/\mathrm{d}x$, and we will use this freely in the future.

2.4 Eigenstates of $\hat{K}$ and the k-representation

Shankar pp 136-137

Recall that we have defined $\hat{K} = -i\hat{D}$ to get a Hermitian operator.

We denote eigenkets of $\hat{K}$ as $|k_0\rangle$, with $\hat{K}|k_0\rangle = k_0|k_0\rangle$ for some specific value k_0, or more generally $\hat{K}|k\rangle = k|k\rangle$ if we don't want to specify the value. Since $\hat{K}$ is Hermitian, allowed values of k_0 must be real.

What is the functional form of $|k\rangle$, $\langle x|k\rangle$? It turns out to be confusing to call this $k(x)$, so we will call it $\phi_k(x)$. In the x-basis, the eigenvalue equation is

$$\langle x|\hat{K}|k\rangle = k\langle x|k\rangle = k\,\phi_k(x)$$

but also, from the x-representation of the operator $\hat{K}$,

$$\langle x|\hat{K}|k\rangle = -i\frac{\mathrm{d}\phi_k}{\mathrm{d}x}$$

Equating the two right-hand sides, we have a familiar differential equation for $\phi_k(x)$, whose solution is

$$\langle x|k\rangle \equiv \phi_k(x) = \sqrt{\tfrac{1}{2\pi}}\,\mathrm{e}^{ikx},$$

where the choice of normalisation will be justified shortly.

Two states of different k must be orthogonal. In fact

$$\langle k|k'\rangle = \int_{-\infty}^{\infty} \langle k|x\rangle\langle x|k'\rangle \,\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-ikx}\mathrm{e}^{ik'x}\,\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{i(k'-k)x}\,\mathrm{d}x = \delta(k-k')$$

and this justifies the choice of normalisation.

This gives us another version of the identity operator

$$\int_{-\infty}^{\infty} |k\rangle\langle k|\,\mathrm{d}k = \hat{I}.$$

By the same argument as used above, for some arbitrary ket $|f\rangle$, $\langle k|f\rangle$ is a function of k, which we will call $F(k)$. Then

$$F(k) \equiv \langle k|f\rangle = \int_{-\infty}^{\infty} \langle k|x\rangle\langle x|f\rangle\,\mathrm{d}x = \sqrt{\tfrac{1}{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-ikx}f(x)\,\mathrm{d}x$$

Thus $F(k)$ is the Fourier transform of $f(x)$. Both are representations of the same abstract ket $|f\rangle$.

Similarly, we can show that

$$f(x) = \sqrt{\tfrac{1}{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{ikx}F(k)\,\mathrm{d}k$$

which is the inverse Fourier transform.

Note that

$$\int_{-\infty}^{\infty} f^*(x)g(x)\,\mathrm{d}x = \langle f|g\rangle = \int_{-\infty}^{\infty} \langle f|k\rangle\langle k|g\rangle\,\mathrm{d}k = \int_{-\infty}^{\infty} F^*(k)G(k)\,\mathrm{d}k$$

which is Parsival's theorem. So if $f(x)$ is square-integrable, so is $F(k)$, and if one is normalised so is the other. The k-basis then maps vectors into an alternative Hilbert space, that of complex square-integrable functions of the real variable k.

We can show that

$$\langle k|\hat{K}|k'\rangle = k\delta(k-k') \qquad \text{and} \qquad \langle k|\hat{X}|k'\rangle = i\frac{\mathrm{d}\delta(k-k')}{\mathrm{d}k}$$

so both operators are local in the k-basis (or k-representation) as well, and we often write $\hat{K} \underset{k}{\longrightarrow} k$ and $\hat{X} \underset{k}{\longrightarrow} i\mathrm{d}/\mathrm{d}k$.

Note that now we have at least three possible representations of $|f\rangle$: as an infinite list of coefficients $\{f_0, f_1, f_2 \ldots\}$ in a basis such as the one introduced at the start, as $f(x)$, or as $F(k)$. All encode the same information about $|f\rangle$, and it is natural to think of $|f\rangle$ as primary, rather than any of the representations.

2.5 Functions in 3-D space

The extension to functions of three coordinates x, y and z is is straightforward. There are operators associated with each, $\hat{X}$, $\hat{Y}$ and $\hat{Z}$, which commute, and corresponding differential operators $\hat{K}_x$, $\hat{K}_y$ and $\hat{K}_z$, which also commute. Between the two sets the the only non-vanishing commutators are $[\hat{X}, \hat{K}_x] = [\hat{Y}, \hat{K}_y] = [\hat{Z}, \hat{K}_z] = i$.

In a more compact notation, we introduce the position operator in 3-d space, $\hat{\mathbf{X}}$, which will be $\hat{X}\mathbf{e}_x + \hat{Y}\mathbf{e}_y + \hat{Z}\mathbf{e}_z$ in a particular coordinate system, and similarly $\hat{\mathbf{K}}$. Boldface now indicates a vector operator, i.e. a triplet of operators. (We have written the 3-D basis vectors $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ instead of $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$.)

The state $|x, y, z\rangle \equiv |\mathbf{r}\rangle$ is an eigenstate of position:

$$\hat{\mathbf{X}}|\mathbf{r}\rangle = \big(\hat{X}\mathbf{e}_x + \hat{Y}\mathbf{e}_y + \hat{Z}\mathbf{e}_z\big)|\mathbf{r}\rangle = \big(x\,\mathbf{e}_x + y\,\mathbf{e}_y + z\,\mathbf{e}_z\big)|\mathbf{r}\rangle = \mathbf{r}|\mathbf{r}\rangle$$
$$\hat{\mathbf{K}}|\mathbf{k}\rangle = \big(\hat{K}_x\,\mathbf{e}_x + \hat{K}_y\,\mathbf{e}_y + \hat{K}_z\,\mathbf{e}_z\big)|\mathbf{k}\rangle = \big(k_x\,\mathbf{e}_x + k_y\,\mathbf{e}_y + k_z\,\mathbf{e}_z\big)|\mathbf{k}\rangle = \mathbf{k}|\mathbf{k}\rangle$$

In position space, $\hat{\mathbf{X}} \longrightarrow \mathbf{r}$ and $\hat{\mathbf{K}} \longrightarrow -i\nabla$.

In 3-D, we have

$$\langle f|g\rangle = \int_{-\infty}^{\infty} f^*(\mathbf{r})g(\mathbf{r})\mathrm{d}^3r; \qquad \langle\mathbf{r}|\mathbf{r}'\rangle = \delta(\mathbf{r}-\mathbf{r}') = \delta(x-x')\delta(y-y')\delta(z-z')$$

The structure of this space is a direct product space: we could write $|x, y, z\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle$, and $\hat{x}_1$ as $\hat{X} \otimes \hat{I} \otimes \hat{I}$. We almost never do, as it is usually not helpful for problems with spherical symmetry. But it enables us to see that the states $\{|m, n, p\rangle\}$ whose wave functions are

$$\langle\mathbf{r}|m, n, p\rangle = H_m(x)\mathrm{e}^{-x^2/2}H_n(y)\mathrm{e}^{-y^2/2}H_p(z)\mathrm{e}^{-z^2/2} = H_m(x)H_n(y)H_p(z)\mathrm{e}^{-r^2/2}$$

are basis functions in the space.

The generalisation of $\phi_k(x)$ is

$$\phi_{\mathbf{k}}(\mathbf{r}) = \langle\mathbf{r}|\mathbf{k}\rangle = \left(\tfrac{1}{2\pi}\right)^{3/2}\,\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}},$$

which is a plane wave travelling in the direction of $\mathbf{k}$.

Caveats

Though we glossed over the fact, the states $|x\rangle$ and $|k\rangle$ do not correspond to functions in the Hilbert space, because they are not square integrable. It is particularly easy to see that $\phi_k(x)$, which is a plane wave of unit magnitude everywhere, is not normalisable, and both $\langle k|k\rangle$ and $\langle x|x\rangle$ are infinite. The x- and k-representations, though, are still extremely useful because they allow us to associate functions and their Fourier transforms with abstract vectors and vice versa. The identity operators are particularly useful for this purpose.

In physical applications the most usual solution to this problem is to imagine the system of interest is in a large box, and require the functions either to vanish at the boundaries, or to be periodic. Then of course only discrete values of the wave vector $\mathbf{k}$ are allowed, but these will be so finely spaced that sums over allowed values can be replaced by integrals, and any dependence on the size of the box drops out. The density of states in statistical physics uses these ideas.

A proper mathematical treatment of functional spaces is well beyond the scope of this course. Griffiths, chapter 3, says a little more about which results of finite-dimensional vector spaces can safely be carried over to infinite-dimensional ones.