Clarifications for 6/30/2011

Keith Fratus

July 6, 2011

Today I talked about the way in which we can represent operators and vectors in terms of a basis. There were two points I wanted to clarify a little bit, which may lead to some confusion. I also wanted to mention something which may help for the homework, regarding Hermitian operators.

1 Representing operators in a basis

One of the ideas I mentioned was the derivative operator acting on the space of second-order polynomials, and its representation in a given basis. The basis vectors I chose were

$$\{|i\rangle\} \quad = \quad \left\{1, x, x^2\right\}, \tag{1}$$

and the way that I chose to represent them was

$$1 \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \; ; \; x \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \; ; \; x^2 \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2}$$

Note that with the usual definition of adding two functions together, or multiplying a function by a number, we can show this space is indeed a vector space. Note that I need not choose an inner product, and notice that without doing so, there is no notion of orthogonality, and, unless I define a norm in some other way without using the inner product, also no notion of normality. The choice of basis above is referred to as a *standard basis*, which is a basis where the n^{th} vector is represented by a column of zeros, except for a 1 for the n^{th} entry. I also chose to introduce the derivative operator, defined by

$$\hat{D}|f\rangle \mapsto |\ \frac{\partial f}{\partial x}\ \rangle, \tag{3}$$

which is to say that the derivative operator maps a function to (not surprisingly) its derivative. How do we represent this operator in the basis I chose? Well, in class, starting with the expression for a vector represented in terms of its components in an orthonormal basis, I derived the result

$$D_{ij} = \langle i|\hat{D}|j\rangle \tag{4}$$

for when we have an inner product. However, I haven't really defined an inner product on this space yet. Despite this though, there's no reason that we need an inner product to define linear maps on a vector space. So there should be a way to do this without using an inner product. And there is. Notice that if I denote the matrix representation of this operator as an arbitrary matrix,

$$\hat{D} \mapsto \begin{pmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{pmatrix}, \tag{5}$$

then the action of the operator on the first basis vector can be represented by,

$$\hat{D}|1\rangle \mapsto \begin{pmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} D_{11} \\ D_{21} \\ D_{31} \end{pmatrix}. \tag{6}$$

While this looks relatively trivial, it is of profound importance. Notice that what we've arrived at is the statement that *the first column of this matrix is the result of acting the operator on the* *first column vector*. Likewise, we could repeat this process with the n^{th} basis vector to get the n^{th} column of the matrix.

To see this idea in action for the derivative operator, notice that because we have,

$$\frac{\partial}{\partial x}(1) = 0; \frac{\partial}{\partial x}(x) = 1; \frac{\partial}{\partial x}\left(x^2\right) = 2x; \tag{7}$$

then we want to assign the correspondence,

$$\hat{D}\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix};\ \hat{D}\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix};\ \hat{D}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}0\\2\\0\end{pmatrix}. \tag{8}$$

Because we know the action of the operator on all n basis vectors, we know what all n of the columns of the matrix are, and so we can immediately write,

$$\hat{D} \mapsto \begin{pmatrix}0 & 1 & 0\\0 & 0 & 2\\0 & 0 & 0\end{pmatrix}. \tag{9}$$

Let's try to use this. We know that,

$$\frac{\partial}{\partial x}\left(5 + 3x + 8x^2\right) = 3 + 16x, \tag{10}$$

while also,

$$\begin{pmatrix}0 & 1 & 0\\0 & 0 & 2\\0 & 0 & 0\end{pmatrix}\begin{pmatrix}5\\3\\8\end{pmatrix} = \begin{pmatrix}3\\16\\0\end{pmatrix}. \tag{11}$$

Calculus with matrices!

Even though I didn't define an inner product on this space, notice that I wrote something interesting in class today on the board, which looked something like

$$D_{12} = \langle 1|\hat{D}|2\rangle = \begin{pmatrix}1 & 0 & 0\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} = 1, \tag{12}$$

and so on for the other entries. This gives the correct result, but it looks a lot like an inner product, which I told you I hadn't defined yet. Can you see why this works from a mathematical standpoint, even though we don't really have an inner product yet? Would it still work if I hadn't used a standard basis? Could I use this idea to then *define* an inner product on this space? I'll let you think about these ideas. Also, notice that while we often define an inner product in terms of some sort of formula, or in terms of some other principle, we don't have to do so; it's always possible to define the inner product by specifying its value for every possible case. For example, I could just order by decree that

$$\langle i|j\rangle \equiv \delta_{ij}, \tag{13}$$

and take this as the definition. Note that this information alone is enough to determine the inner product for any other two vectors, since these vectors can be written in terms of the basis, and then using the linearity of the inner product, we can compute this inner product in terms of the inner products of the basis vectors.

This means of defining the inner product happens quite often in physical applications in quantum mechanics. Often times we will take the physical states of a system to be the "basis vectors," and then merely demand that they be orthonormal.

2 Vectors and their representations

The other idea I talked about was the difference between a vector and its representation in a basis. The idea I tried to drive home was that the physical or abstract vector itself is not the same as the column of numbers which represents it in a basis, much the same way that the physical object which is my house is not the exact same thing as the number which is its street address. But it turns out I lied just a little bit.

While in almost every situation we'll be confronted with, the column of numbers associated with a vector is not the vector itself, if we really wanted to, we could actually choose to make this be the case. Just in the way that we could study matrices as the vectors themselves, if we really insisted, we could choose to have a vector space of n-tuples, where *the literal columns of numbers are the vectors themselves.* We already showed in class that with a suitable definition of addition and scalar multiplication, these objects form a vector space, so there is nothing wrong with this. But things can get subtle when we talk about representations of these vectors.

Let's choose a particular basis, for example

$$|1\rangle \mapsto \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix};\ |2\rangle \mapsto \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}. \tag{14}$$

We can even define an inner product, which, for two vectors

$$|a\rangle = \begin{pmatrix}a_1\\a_2\end{pmatrix};\ |b\rangle = \begin{pmatrix}b_1\\b_2\end{pmatrix}, \tag{15}$$

we will take to be

$$\langle a|b\rangle \equiv a_1 b_1 + a_2 b_2, \tag{16}$$

where for now I'm taking these vectors to be real-valued (of course technically speaking I still need to verify this operation obeys all of the required properties of an inner product for me to be able to call it one!). Notice that the above columns are *not* representations. I'm taking these columns of numbers to be the vectors themselves, which is why I've used equal signs, and not correspondence arrows (can you show that the basis I chose is orthonormal under this inner product?).

Now consider the vector

$$|c\rangle = \begin{pmatrix}1\\0\end{pmatrix}. \tag{17}$$

Let's write this in terms of our earlier basis, which would look like,

$$|c\rangle = \begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix} = \frac{1}{\sqrt{2}}\left(|1\rangle + |2\rangle\right). \tag{18}$$

Please think about this result carefully. What we've just decided for ourselves is that their are *two* columns of numbers associated with our vector of interest, but they have dramatically different interpretations. The column of numbers given by

$$\begin{pmatrix}1\\0\end{pmatrix} \tag{19}$$

is *the vector itself*, while the column of numbers

$$\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix} \tag{20}$$

is *its representation in a particular basis.* The first column of numbers will always be the same, because that is the vector itself, but the second column of numbers will depend on our choice of basis. We can see that whenever we take *mathematical entities themselves to be vectors*, we can get ourselves seriously confused if we're not careful to distinguish these two ideas!

Notice that there are two ways to calculate the square of the norm of this vector. We can use the original definition, as a formal operation on the vector itself,

$$a_1 a_1 + a_2 a_2 = 1 * 1 + 0 * 0 = 1, \tag{21}$$

or, I could calculate it using its coefficients in the orthonormal basis,

$$\langle c|c\rangle = \alpha_1\alpha_1 + \alpha_2\alpha_2 = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = 1. \tag{22}$$

Remember that if I have

$$|a\rangle = \sum_i \alpha_i |e_i\rangle \ ; \ |b\rangle = \sum_i \beta_i |e_i\rangle, \tag{23}$$

then it is always true that,

$$\langle b|a\rangle = \sum_j \sum_i \beta_j^* \alpha_i \ \langle e_j|e_i\rangle, \tag{24}$$

which for an orthonormal basis reduces to

$$\langle b|a\rangle = \sum_i \beta_i^* \alpha_i. \tag{25}$$

The object

$$g_{ij} \equiv \langle e_i|e_j\rangle \tag{26}$$

is referred to as the metric tensor, or the first fundamental form, and is very fundamental to the study of differential geometry. You'll see it all the time if you take a course in General Relativity.

3 The Adjoint of an Operator and its Representation

One last piece of confusion that seems to be coming up a lot is the idea of an operator adjoint, and its representation in a basis. While we usually think of the adjoint of an operator in terms of swapping its rows and columns in a matrix, and then taking the complex conjugates, this is not really the fundamental definition, especially if we do not choose to write the operator in terms of a matrix at all. For example, for the derivative operator acting on square-integrable functions on the infinite real line, this matrix representation in a basis of Fourier modes would actually have *continuous* indices, and so this "matrix" would look more like a continuous plane than a discrete set of points on a lattice.

The definition of the adjoint is in terms of the dual correspondence. If we have the mapping

$$|b\rangle = \hat{A}|a\rangle, \tag{27}$$

then the adjoint of this operator is defined by

$$\langle b| = \langle a|\hat{A}^\dagger. \tag{28}$$

Keep in mind that operators always act on bras to the left, which is to say that the following inner product can be viewed in two ways, with $\hat{A}^\dagger$ either acting on the bra, or the ket;

$$\langle a|\hat{A}^\dagger|b\rangle = \left(\langle a|\hat{A}^\dagger\right)|b\rangle = \langle a|\left(\hat{A}^\dagger|b\rangle\right). \tag{29}$$

Another notation which is also used is

$$\langle \hat{A}a| \equiv \langle a|\hat{A}^\dagger, \tag{30}$$

which conveys the idea of first acting the operator on a ket, and then mapping the result to the dual.

Of course, when doing actual calculations, we often want to use this definition to find the components of the operator in a basis. Notice that

$$\hat{A}^\dagger_{ij} = \langle e_i|\hat{A}^\dagger|e_j\rangle = \left(\langle e_i|\hat{A}^\dagger\right)|e_j\rangle \tag{31}$$

in the basis $\{|e_i\rangle\}$. Now, we know for any two vectors,

$$\langle b|a\rangle = \langle a|b\rangle^*, \tag{32}$$

and so if we take,

$$\langle b| = \langle e_i|\hat{A}^\dagger \; ; \; |b\rangle = \hat{A}|e_i\rangle, \tag{33}$$

where the second equality follows from the definition of the adjoint, along with,

$$|a\rangle = |e_j\rangle, \tag{34}$$

then we get,

$$A^\dagger_{ij} = \langle e_i|\hat{A}^\dagger|e_j\rangle = \left(\langle e_j|\hat{A}|e_i\rangle\right)^* = A^*_{ji} \tag{35}$$

which is the usual expression we are more familiar with.

However, for problem one on the second homework, it will generally be easier to work with another statement of Hermiticity. When I derived the statement

$$\langle e_i|\hat{A}^\dagger|e_j\rangle = \left(\langle e_j|\hat{A}|e_i\rangle\right)^*, \tag{36}$$

there was no reason I had to pick basis vectors in the inner product, and in general, we should have for all vectors

$$\langle a|\hat{A}^\dagger|b\rangle = \left(\langle b|\hat{A}|a\rangle\right)^*, \tag{37}$$

and if $\hat{A}$ is Hermitian, so that it equals its own adjoint, we arrive at the conclusion,

$$\langle a|\hat{A}|b\rangle = \left(\langle b|\hat{A}|a\rangle\right)^*, \tag{38}$$

if and only if $\hat{A}$ is Hermitian. In a lot of situations, it is easier to verify Hermiticity using this expression. Really, hidden behind this is the same formula we're all used to, since if the above relation holds for basis vectors (which gives the relation for the components in a basis), it will hold for all vectors (can you show this?)

Again, keeping in mind the difference between an operator adjoint and its representation in a basis can help avoid a lot of confusion.