Physics 342 Lecture 2

Welcome back to class on this lovely Wednesday morning! Last lecture, we had ended with an intriguing observation regariding the Taylor expansion as an "exponentiation" of the derivative. That is, we identified:

$$f(x+\Delta x) = \sum_{n=0}^{\infty} \frac{\Delta x^n}{n!} \frac{\partial^n}{\partial x^n} f(x) \equiv e^{\Delta x \frac{\partial}{\partial x}} f(x).$$

The odd exponential of a derivative is <u>defined</u> through its Taylor expansion. In this derivation/construction, we exploited the <u>linearity</u> of the derivative as an operator. In this lecture, we are going to first remind about properties of linear operators and then use that to provide a profound new interpretation of the derivative.

You are most likely familiar with the property of <u>linearity</u> from a course on <u>linear</u> algebra (the names are not a coincidence...). For a matrix $\mathbb{M}$ and vectors $\vec{v}$ and $\vec{u}$, the property of linearity means that

1) $\mathbb{M}(\vec{v}+\vec{u}) = \mathbb{M}\vec{v} + \mathbb{M}\vec{u}$ and

2) $\mathbb{M}(a\vec{v}) = a\mathbb{M}\vec{v}$, where a is just a number (a <u>scalar</u>).

We can generalize this definition of linearity to define a linear operator $\hat{\mathcal{O}}$ as follows. For two functions f and g, a linear operator satisfies:

1) $\hat{\mathcal{O}}(f+g) = \hat{\mathcal{O}}f + \hat{\mathcal{O}}g$ and

2) $\hat{\mathcal{O}}(af) = a\hat{\mathcal{O}}f$, where a is a scalar number.

Many familiar objects are linear operators. For example, a simple function of position is a linear operator. If $\theta \equiv \theta(x)$ is just a function of x, then indeed it satisfies:

$$1)\quad \theta(x)\Big(f(x)+g(x)\Big) = \theta(x)f(x) + \theta(x)g(x)$$

$$2)\quad \theta(x)\Big(a f(x)\Big) = a\,\theta(x) f(x),$$

by the distributive law of multiplication and commutivity of multiplication of numbers. Connecting with our study from last lecture, the derivative $\partial/\partial x = \theta$ is also a linear operator. That is, for two functions $f(x)$ and $g(x)$ and scalar number a, we have:

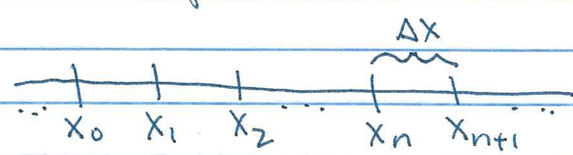
$$1)\quad \frac{\partial}{\partial x}\Big(f(x)+g(x)\Big) = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}$$

$$2)\quad \frac{\partial}{\partial x}\Big(a f(x)\Big) = a\frac{\partial f}{\partial x}.$$

We will encounter _many_ more linear operators throughout this course, but for now, I will just focus on the derivative as our "canonical" linear operator.

I think you have a good sence and intuition for matrices: They act on vectors to rotate or scale them, they have a set of eigenvectors and eigenvalues that encode intrisic information about the action and rank of the matrix, there are algorithmic procedures for calculating the determinant, etc. At its core, however, all a matrix is is a linear operator that acts on vectors, and _all_ of these properties of a matrix that I just mentioned simply follow from its linearity. So, if this is the case, I should be able to think about _any_ linear operator as a matrix with concrete elements at a given row and column location. If this is the case, then the derivative, as a linear operator, somehow can be thought of as a matrix. However, the derivative acts on functions $f(x)$, not vectors, and there's no sense in which you can identify the element of the derivative in the, say, third row, fifth column. Or can you?

To provide a "matrix" interpretation of the derivative, we will think of our space in x as discrete. That is, we will put the x-coordinate on a grid:



where the grid spacing is Δx. As $\Delta x \to 0$, we recover the continuous real-x line. On this grid, there is "nothing" between neighboring points, so we have to modify on definition of the derivative appropriately. First, for some function $f(x)$, on this grid, it is only evaluated at discrete x, so we denote

$f(x_n) \equiv f_n$, the value of f at the grid point n.

Then, recalling the very first idea of a derivative from intro calculus, the derivative of f at gridpoint n can be defined as:

$$\frac{\partial f_n}{\partial x} \equiv \frac{f(x_{n+1}) - f(x_{n-1})}{2\Delta x} = \frac{f_{n+1} - f_{n-1}}{2\Delta x}.$$

As $\Delta x \to 0$, this of course returns the standard, continuous derivative we know and love. However, this form makes the matrix-property of the derivative clear. Let's construct a vector of function values on the grid:

$$\vec{f} = \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \\ \vdots \end{pmatrix}$$

This vector may be infinite in length but let's not worry about that. It is atleast not continuous.

Now the action of the derivative $\partial/\partial x$ on this grid is simply as a matrix that multiplies $\vec{f}$:

$$\frac{\partial f}{\partial x} \to \mathbb{D}\vec{f} = \begin{pmatrix} \ddots & & & & \\ -\frac{1}{2\Delta x} & 0 & \frac{1}{2\Delta x} & 0 & \cdots \\ & -\frac{1}{2\Delta x} & 0 & \frac{1}{2\Delta x} & 0 \ \cdots \\ & 0 & -\frac{1}{2\Delta x} & 0 & \frac{1}{2\Delta x} \ \cdots \\ & 0 & 0 & 0 & 0 \ \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ f_{n-1} \\ f_n \\ f_{n+1} \\ \vdots \end{pmatrix}$$

This matrix has 0s on the digonal, and the entries immediately off the diagonal are $\pm\frac{1}{2\Delta x}$, depending on if the point is to the left or right of the point of interest. Clearly, this matrix ceases being sensible as $\Delta x \to 0$, as the off diagonal elements diverge. However, this picture of a linear operator, any linear operator, as a matrix will be central to our language in quantum mechanics.

Another thing that is useful when thinking about matrices is the value of their individual elements. Consider a matrix $\mathbb{M}$, and let M_{ij} be the value (i.e. "matrix element") at row i and column j. How do we identify such an element? Well just given $\mathbb{M}$, the way we do it is to sandwhich the matrix between two vectors that only have non-zero entries at the i^{th} and j^{th} locations. That is, for vectors

$$\vec{v}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{th} \text{ entry} \qquad \vec{v}_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j^{th} \text{ entry}$$

the matrix element M_{ij} is

$$\vec{v}_i^{\,T} M \vec{v}_j = M_{ij}.$$

For example, for the 2x2 matrix $M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$ note that:

$$\vec{v}_1^{\,T} M \vec{v}_2 = (1\ 0) \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = M_{12}, \text{ as promised.}$$

Now, it is useful to be a bit more explicit with the dot products multiplication to illustrate what is going on. Note that we can express matrix multiplication as:

$$\vec{v}_i^{\,T} M \vec{v}_j = \sum_{a,b} (v_i)_a M_{ab} (v_j)_b = \sum_{a,b} \delta_{ia} M_{ab} \delta_{jb} = M_{ij}.$$

There are a couple things going on here. First, we denote the entries of the matrix and vector by the indices a and b. Matrix multiplication means to sum over a and b the product of the elements of the vectors and the matrix. Next, we note that the vectors only have non-zero entries when $a = i$ (for $\vec{v}_i$) and $b = j$ (for $\vec{v}_j$). So, we can replace the vector elements with Kronecker δ's:

$$\delta_{ab} = \begin{cases} 0 & \text{if } a \neq b \\ 1 & \text{if } a = b. \end{cases}$$

The Kronecker δ's then pick out the matrix element M_{ij}.

Further, note the important property that the length of the two vectors is each 1:

$$\vec{v}_i \cdot \vec{v}_i = 1, \quad \vec{v}_j \cdot \vec{v}_j = 1 \Rightarrow \vec{v}_i \cdot \vec{v}_j = \delta_{ij}.$$

This is necessary so that the product $\vec{v}_i^{\,T} \mathbb{M} \vec{v}_j$ returns M_{ij}, and not M_{ij} scaled by some factor.

So, with this understanding, let's go back to the derivative as a matrix and extract its elements. The $(i,j)^{th}$ element of the derivative on a grid $\mathbb{D}$ is:

$\vec{f}_i^{\,T} \mathbb{D} \vec{f}_j = D_{ij}$, as a property of matrix multiplication.

For this to hold, we must require that the length of the vectors $\vec{f}_i$ and $\vec{f}_j$ are both 1. Considering $\vec{f}_i$ first this is

$$1 = \vec{f}_i \cdot \vec{f}_i = \vec{f}_i^{\,T} \vec{f}_i = \sum_a (f_i^T)_a (f_i)_a.$$

I've left the sum unevaluated, and also left the transpose on the f_i vector on left. Recall that this function is defined on a grid of space Δx. This normalization must be independent of the grid spacing Δx: the vector squares to 1 for any Δx. The number of grid points in x is proportional to $1/\Delta x$; that is, if we consider the range $x \in [0,5]$, then the number of grid points $N = 5/\Delta x$. To ensure that the sum is independent of Δx, this requires the transpose vector to be:

$$\vec{f}_i^{\,T} \vec{f}_i = \sum_{a=1}^{N} (f_i^T)_a (f_i)_a = \sum_{a=1}^{N} \Delta x (f_i)_a (f_i)_a = \sum_{a=1}^{N} \Delta x (f_i)_a^2.$$

Correspondingly, we have

$$1 = \sum_{a=1}^{N} \Delta x (f_i)_a^2 = \sum_{a=1}^{N} \Delta x (f_j)_a^2 \,; \quad \delta_{ij} = \sum_{a=1}^{N} \Delta x (f_i)_a (f_j)_a .$$

This identification in the discrete case then tells us how to interpret the continuous case. As $\Delta x \to 0$, these dot products just return the definition of the integral:

$$\lim_{\Delta x \to 0} \sum_{a=1}^{N} \Delta x (f_i)_a^2 \to \int_{x_0}^{x_1} dx\, f_i(x)^2 = 1,$$

for some function $f_i(x)$ on the domain $x \in [x_0, x_1]$. This requirement on a function is called an "L^2-norm". It is the generalization of the Pythagorean|Euclidean norm to continuous functions. Now, we have a prescription to determine the $(i,j)^{th}$ entry of the real, continuous, derivative. For two (appropriate...) functions $f_i(x)$, $f_j(x)$ that are L^2-normalized, the $(i,j)^{th}$ entry of the linear operator $\partial/\partial x$ is:

$$\left(\frac{\partial}{\partial x}\right)_{ij} = \int_{x_0}^{x_1} dx\, f_i(x) \frac{\partial}{\partial x} f_j(x), \quad \text{on } x \in [x_0, x_1].$$

Fascinating... To end today, I want to note that the individual entries of a matrix depend on the basis|coordinate system that you choose. So, in some well-defined sense, individual matrix elements are meaningless, because we know that everyone only agrees upon basis-independent quantities. What are these "basis-independent" objects and how can we exploit them? More next time...