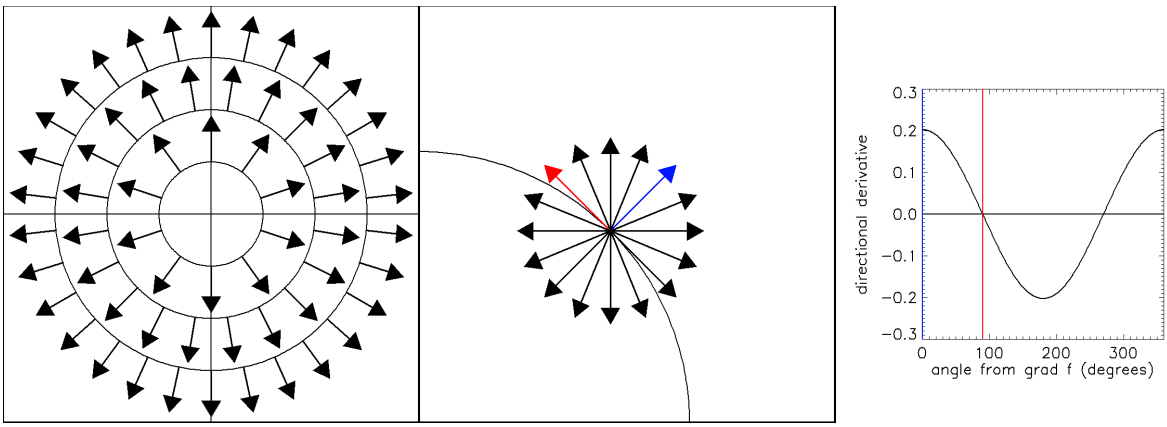
Math for Electromagnetism



©Eric Donovan

Introduction

At this point you have seen and worked with Maxwell's Equations in integral form:

$$\int \vec{E} \cdot \vec{da} = \frac{Q_{enclosed}}{\epsilon_\circ} \qquad \oint \vec{E} \cdot \vec{dl} = -\frac{d\Phi_M}{dt} \qquad \int \vec{B} \cdot \vec{da} = 0 \qquad \oint \vec{B} \cdot \vec{dl} = \mu_\circ I_{enclosed} + \mu_\circ \epsilon_\circ \frac{d\Phi_E}{dt}$$

I know that you have some tools to deal with these equations. I also know that the concepts and math tools that enable their use are relatively new to most if not all of you. Moreover, we will soon be dealing with these and other integral equations in their differential forms. Further, we will be using specific tools from calculus and vector calculus, right from the start. In fact, although this is not a mathematics course, the material we cover throughout is fairly mathematical.

This is my attempt to provide a moderately high-level but still comprehensive view of the mathematics necessary for Electro*dynamics* at the senior undergraduate level. This is not a math course, but I believe that it is good to have a gut-level, if not rigorous, understanding of how these tools are developed from more fundamental concepts. Rather than derive, I attempt to provide ways of visualizing intuitively why these things work the way they do. An example is that by understanding the definition of divergence, you can easily see why the Divergence theorem has to be true. This falls short of *proof*, but in this sequence of courses, the goal is to be able to use these concepts.

For me, the math course where I learned most of these concepts was 32 years ago. So I can go back and work through the formal proofs, but usually I just use the concepts. However, I feel better about using them when I can convince myself I know roughly how whatever it is that I am using is derived. I present this in a quasi-text-like manner, meaning I explain some things, and ask you to work through some things. I believe working through this material, start to finish, will be very helpful for you with this course and beyond. This is my first attempt at putting this together, so please understand this is a "work in progress". I want constructive feedback, and I will fix and add to this on the course web page as we go.

Some Basics

Expressed as questions in the Question Bank, these are basic concepts you should be absolutely clear on.

The Dot Product

The *dot* between two vectors is defined as

$$\vec{A} \cdot \vec{B} = AB\cos(\theta) \tag{1}$$

where A and B are the magnitudes of $\vec{A}$ and $\vec{B}$, respectively, and θ is the smallest angle between them.

Of course, you already know the dot product, and more than that the fact that the dot product is *also* given by (in Cartesian coordinates)

$$\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z \tag{2}$$

Now, we are not free to simply define the dot product in two different ways. We have to pick one of Equation 1 or Equation 2 to be our definition, and then show the other follows. It is left as an exercise to show that, at least in 2D, 2 follows from 1.

The Cross Product

Let us define the *cross product* between two vectors to be

$$\vec{A} \times \vec{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} = (A_y B_z - A_z B_y)\hat{x} + (A_z B_x - A_x B_z)\hat{y} + (A_x B_y - A_y B_x)\hat{z} \tag{3}$$

Of course there is another representation we all know about, namely

$$\vec{A} \times \vec{B} = AB\sin(\theta)\hat{n} \tag{4}$$

where A and B are the magnitudes of $\vec{A}$ and $\vec{B}$, respectively, θ is the smallest angle between them, and $\hat{n}$ is a unit vector orthogonal to both $\vec{A}$ and $\vec{B}$, with the direction chosen via the right hand rule. It is left as an exercise to demonstrate *the equivalence of the two representations*, namely that Equation 4 follows from Equation 3

I want to remind you, the reader, and presumably the student, what my goal is. I want you to master the physics, and be able to use the required math in the process.

There are (at least) two representations of the dot product

$$\vec{A} \cdot \vec{B} = AB\cos(\theta) \quad \text{and} \quad \vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z$$

and (at least) two representations of the cross product

$$\vec{A} \times \vec{B} = AB\sin(\theta)\hat{n} \quad \text{and} \quad \vec{A} \times \vec{B} = \Big[A_y B_z - A_z B_y\Big]\hat{x} + \Big[A_z B_x - A_x B_z\Big]\hat{y} + \Big[A_x B_y - A_y B_x\Big]\hat{z}$$

On a test, in a senior E&M class, I would never ask you to reconcile the two (in either case). Rather, I want you to know they are equivalent, and you can show they are equivalent. Plus, I want you to be able to use either, in each case.

Derivatives

The derivative is the measure of the change of a function due to an infinitessimal change in a variable upon which it depends. Every continuous function f of a single variable looks like a line if looked at on a small enough range Δx. Thus, a function can be thought of as having a *slope*

$$m = \frac{\Delta y}{\Delta x}$$

In the limit of small Δx, the slope approaches the slope of the line tangent to the curve at that point. In the limit of small Δx, this slope is the rate of change of the function with x (or t or whatever), which we call the derivative:

$$\frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \tag{5}$$

where I adopt the more common f and h notation.

This applies to functions of multiple variables, wherein we take the same limit as above, while keeping the other variables constant. Imagine we have $f(x, y, z, t)$, then we can ask ourselves how f changes with x while holding y, z, and t constant. With the three other independent variables held constant, we are again looking at the slope of a curve, and the *partial* derivative is

$$\frac{\partial f}{\partial x} = \lim_{h\to 0} \frac{f(x+h, y, z, t) - f(x, y, z, t)}{h}$$

and the corresponding partial derivatives with respect to y, z, and t:

$$\frac{\partial f}{\partial y} = \lim_{h\to 0} \frac{f(x, y+h, z, t) - f(x, y, z, t)}{h},$$

$$\frac{\partial f}{\partial z} = \lim_{h\to 0} \frac{f(x, y, z+h, t) - f(x, y, z, t)}{h},$$

and

$$\frac{\partial f}{\partial t} = \lim_{h\to 0} \frac{f(x, y, z, t+h) - f(x, y, z, t)}{h}$$

Integrals

Integration was developed by Isaac Newton to address a specific set of physics questions including gravitation. I'm always amazed by integration, because it's a sort of sloppy finite mathematical concept, where in the limit of large numbers of finite elements, all the sloppiness becomes completely immaterial. As I repeatedly state, this is not a math course. So I am going to go through this material in a way that supports the use of the concept for physics, and in particular, electrodynamics.

Imagine we have a function $f(x)$. You know that the quantity

$$\int_{x_0}^{x_1} f(x)dx$$

is the area under the curve between x_0 and x_1. Formally, the area is approximated by a sum, the limit of which is *exactly* the area, and hence the integral

$$\int_{x_0}^{x_1} f(x)dx = \lim_{\substack{\Delta x_i \to 0 \\ N\to\infty}} \sum_{i=0}^{N} \tilde{f}_i \Delta x_i$$

But this limit has a few caveats. First, $\tilde{f}_i$ is any value of $f(x)$ on the interval Δx_i, meaning it could be the midpoint, the average value of $f(x)$ on that interval, the value of $f(x)$ at a randomly chosen point on Δx_i, ... really, the value of $f(x)$ at any point on the interval. Second, the intervals do not need to be of uniform length. Third, the limit, which when I was in first year calculus at the University of Western Ontario was called "the mesh limit", is for the number of elements going to ∞ while at the same time the maximum value of every Δx_i is going to zero. In Figure 1, I show a function between x_0 and x_1 which is broken up into five intervals (Δx_i). In this case, the area under the curve between x_0 and x_1 is estimated by sum of the areas of the the five rectangles as shown. Imagine increasing the number of rectangles, and decreasing their maximum width. Eventually, how one picks $\tilde{f}_i$ for each interval becomes irrelevant.

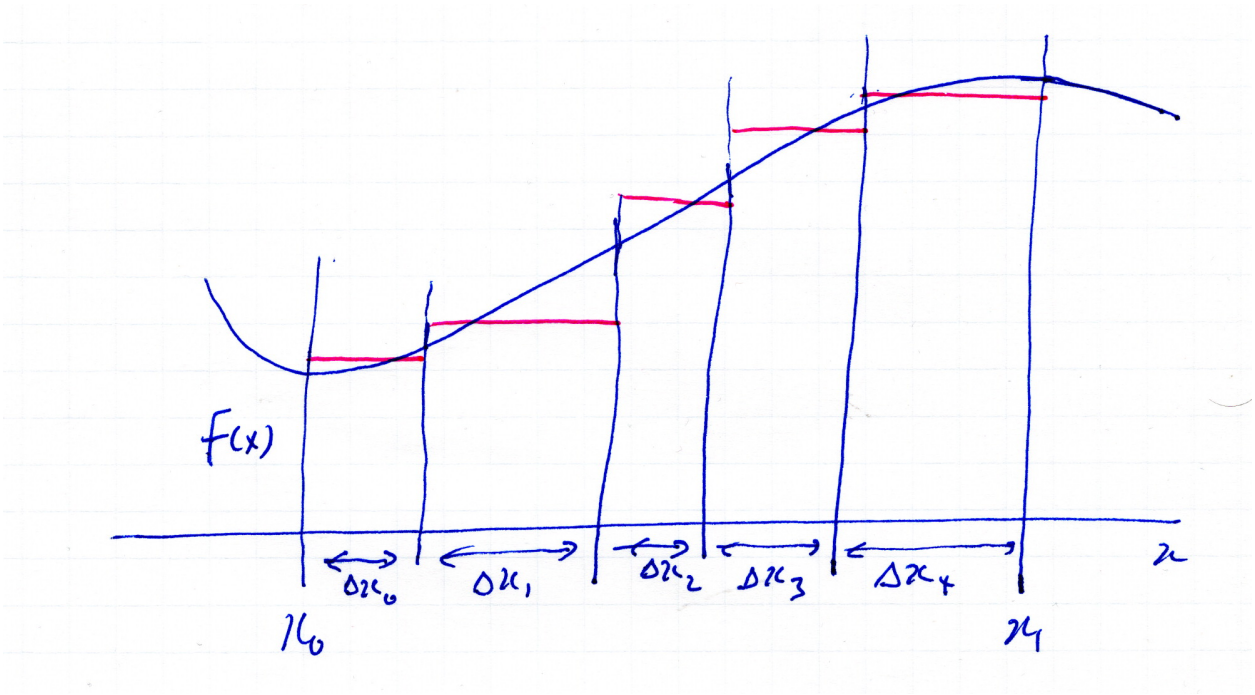


Figure 1: Simple heuristic picture of how atoms respond to an applied electric field in dielectrics.

The First Fundamental Theorem of Calculus is easy to understand within this context. If you define the derivative to be

$$\frac{dF(x)}{dx} = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h}$$

and

$$F(x) = \int_{x_0}^{x} f(x')dx'$$

then the derivative of $F(x)$ is

$$\frac{dF(x)}{dx} = \lim_{h\to 0} \frac{1}{h} \int_{x}^{x+h} f(x')dx'$$

which is the area under the curve ($f(x)$) between x and $x+h$ (see Figure 2).

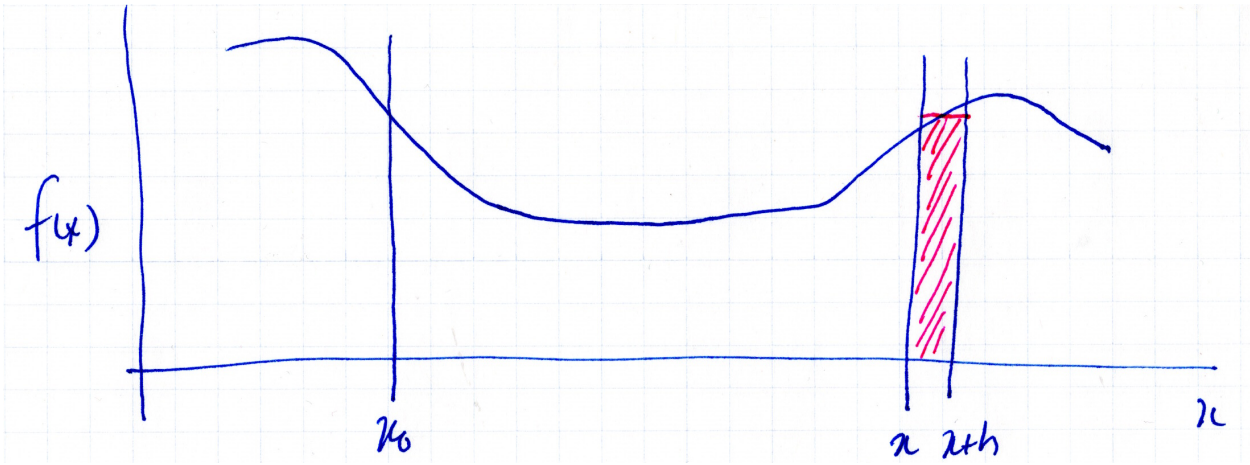


Figure 2: Geometric basis for the First Fundamental Theorem.

For small h, this area (see Figure 2) approaches $hf(x)$, so that we have

$$F(x) = \int_{x_0}^{x} f(x')dx' \quad \text{and} \quad \frac{dF(x)}{dx} = f(x)$$

We should understand integration as summation, the adding up of many little bits of something (in this case area elements). We also see that differentiation and integration undo each other, each being the other's 'anti-operation.' Finally, I am treating the function $f(x)$ as though it is *well behaved*, meaning it is continuous, its derivative is continuous, etc. Of course there are many situations where this is not true, but in each case we will either deal with a problem *piecewise* or use some mathematical subtlety to deal with, for example, a singularity.

Path Integrals

There are many applications in physics that require carrying out integrals *along* (or *on*) a path or contour. These are called contour, path, or line integrals. The names are interchangeable with the caveat that *usually* contour integral is associated with a path integral in the complex plane. However, there is no hard and fast rule, and I prefer the term contour integral in general for two reasons. First, it's what they were called in my program as an undergraduate. Second, and more importantly, many of the integrals we do are along contours that bound surfaces, and I find "path integral along the contour" cumbersome and actually kind of odd (I mean the integral in Stokes' Theorem in the next section is "*along the contour that bounds the surface S*". Regardless, I will use 'contour-', 'path-' and 'line-integral' interchangeably in this material.

Integrals as Averages

When we worked through the First Fundamental Theorem, above, we ended up with

$$F(x) = \int_{x_\circ}^{x} f(x')dx' \quad \text{and} \quad \frac{dF}{dx} = \lim_{h\to 0} \frac{1}{2h} \int_{x_\circ - h}^{x_\circ + h} f(x')dx' = f(x)$$

In doing so, the last equality, we appealed to the fact that if the function is well behaved, then for a diminishingly small interval, the integral is the value of the function at x times the length of the interval. This is in one sense obviously true, however it is not exactly rigorous.

A better approach is to consider the relationship between the average of a function on an interval and the integral of that function on that interval. Looking at the interval $(x-h, x+h)$, the average of the function $f(x)$ on that interval is *defined* to be

$$\int_{x-h}^{x+h} f(x')dx' = 2hf^*$$

where I use f^* to represent the average. This is shown in Figure 3. The average value of the function is shown by the dashed red line. The area under the red dashed line equals that under the curve $f(x)$ between $x_\circ - h$ to $x_\circ + h$.

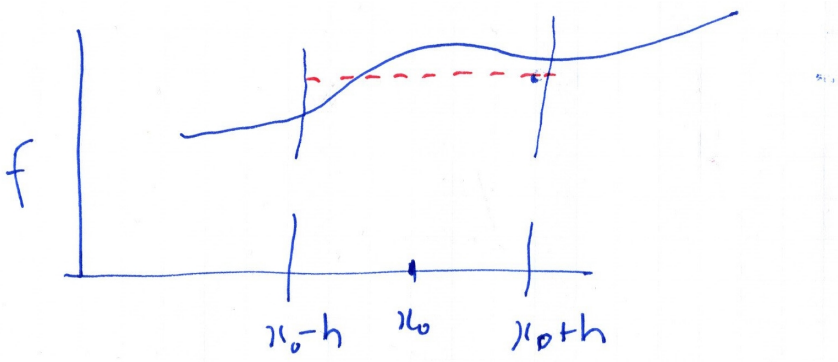


Figure 3: Between $x_\circ - h$ to $x_\circ + h$ the area under the curve $f(x)$ is equal to the area under $f(x)$ between $x_\circ - h$ to $x_\circ + h$ (that is the definition of *average* in a mathematical sense).

You can (and sometimes should) represent integrals in any dimensional space in terms of averages. In 1D,

$$\int_{x_\circ}^{x_\circ + \Delta x} f(x)dx = \Delta x f^* \tag{6}$$

In 2D,

$$\int_S f da = Af^* \quad \text{where} \quad A = \int_s da$$

And in 3D,

$$\int_V f d\tau = Vf^* \quad \text{where} \quad V = \int_V d\tau$$

There is a corollary. Looking at Equation 6, and imagining the function is well behaved, then for smaller intervals we expect the average of the function to be closer and closer to its value everywhere on the interval. Consequently,

$$\lim_{\Delta x \to 0} f^* = f(x)$$

With this, Equation 6 gives

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_x^{x+\Delta x} f(x')dx' = \lim_{\Delta x \to 0} \frac{f^* \Delta x}{\Delta x} = \lim_{\Delta x \to 0} f^* = f(x)$$

or, simply,

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_x^{x+\Delta x} f(x')dx' = f(x) \tag{7}$$

As an aside, convince yourself that all three of these limits yield the same result

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_x^{x+\Delta x} f(x')dx' = f(x), \quad \lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_{x-\Delta x}^{x} f(x')dx' = f(x) \text{ and } \lim_{\Delta x \to 0} \frac{1}{2\Delta x} \int_{x-\Delta x}^{x+\Delta x} f(x')dx' = f(x) \tag{8}$$

Analogously, for any family of circular surfaces of radius R centered on the point $\vec{r}$,

$$\lim_{R \to 0} \frac{1}{A} \int_S f(\vec{r}')da' = f(\vec{r}) \tag{9}$$

Similarly, for a family of spherical volumes of radius R centered on the point $\vec{r}$,

$$\lim_{R \to 0} \frac{1}{V} \int_V f(\vec{r}')d\tau' = f(\vec{r}) \tag{10}$$

These results, though at once obvious and obscure, will prove remarkably useful as we meander through electrodynamics. A final note, a consequence of the Mean Value Theorem, is that we can actually be quite sloppy about this. Provided the functions $f(x)$ is continuous on the interval for the definite integral, it is *exactly* true that

$$\frac{1}{x_1 - x_0} \int_{x_0}^{x_1} f(x)dx = f(x^*)$$

where x^* is *at least* one value of x between x_0 and x_1. A different way of expressing this is that *a continuous function assumes its average value at at least one point on an interval.* This 1D concept obviously extends to 2D and 3D.

Equality of Integral Quantities

Imagine we know that

$$\int_a^b f(x)dx = 0$$

for all *finite* intervals (a, b). It seems obvious that for this to be true, it *must* be the case that for all x, $f(x) = 0$, but saying it is obvious does not prove it is. We can prove this as follows. We can prove this as follows.

Suppose there is a finite interval (a, b) where $f(x) > 0$. Then its average value on (a, b) is greater than zero. But it's average value on that interval is

$$< f >= \frac{1}{b-a}\int_a^b f(x)dx$$

However, we are given that

$$\int_a^b f(x)dx = 0$$

for all finite finite intervals (a, b). Therefore there is no finite interval where $f(x) > 0$. By the same argument, supposing that there is a finite interval where $f(x) < 0$, on which its average is negative, so there can be no finite interval where $f(x) < 0$. Therefore, there are no finite intervals where $f(x) \neq 0$, therefore $f(x) = 0$ everywhere.

This is true for all functions that are in some sense *well behaved*, which we can count on for any functions we deal with in electromagnetism. Please do not follow this squirrel, but there are functions for which this argument does not hold, such as

$$f(x) = \lim_{n\to\infty} \sin(nx)$$

It is straightforward to show (but please don't take the time at this juncture) that while

$$\int_a^b \left[\lim_{n\to\infty} \sin(nx)\right]dx = 0$$

for all finite intervals (a, b), it is also true that there are an infinite number of points on the $x - axis$ where $f(x)$ is not zero (the limit is taking the distance between adjacent zeros of $\mathrm{Sic}(x)$ and bringing them arbitrarily close to each other, but however close they are the function must go from 0 to ± 1 between those adjacent zeros, because it is after all a sinusoind of amplitude 1). Thus, in this case there are no *finite* intervals where $f(x) \neq 0$. But this is a weird function, a mmber of the exotic functions including the Dirac Delta Function, that mathematicians call *generalized functions*. In this case,

$$f(x) = \lim_{n\to\infty} \sin(nx)$$

is said to be *zero in the sense of generalized functions*, because

$$\int_a^b \left[\lim_{n\to\infty} \sin(nx)\right]g(x)dx = 0$$

for all *well behaved functions* $g(x)$, and all intervals (a, b). But I digress!

In 1D, we have the following. If

$$\int_a^b f(x)dx = 0$$

for all finite intervals (a, b), then $f(x) = 0$ everywhere.

It is completely straightforward to extend this concept to two dimensions, so we can say that if

$$\int_A f da = 0$$

for all finite surfaces A, then $f = 0$ everywhere.

Similarly, in three dimensions, we can say that if

$$\int_V f d\tau = 0$$

for all finite volumes V, then $f = 0$ everywhere.

We use this very simple, intuitive, and very powerful concept repeatedly in electromagnetism at this level. I'm going to summarise these results which, going forward, we will use as em rules, for lack of a better term. As well, for simplicity, I drop the word *discrete* on th understanding that these rules apply to *well behaved* functions, which we can count on functions to be when we are dealing with physics at this level. To summarize:

$$\int_a^b f(x)dx = 0 \quad \text{for all intervals } (a, b) \text{ means } f(x) = 0, \text{ everywhere.}$$

$$\int_A f da = 0 \quad \text{for all areas } A \text{ means } f = 0, \text{ everywhere.}$$

$$\int_V f d\tau = 0 \quad \text{for all volumes } V \text{ means } f = 0, \text{ everywhere.}$$

Obvious corollaries to each of these, that we can also use as rules, are

$$\int_a^b f(x)dx = \int_a^b g(x)dx \quad \text{for all intervals } (a, b) \text{ means } f(x) = g(x), \text{ everywhere.}$$

$$\int_A f da = \int_A g da \quad \text{for all areas } A \text{ means } f = g, \text{ everywhere.}$$

$$\int_V f d\tau = \int_V g d\tau \quad \text{for all volumes } V \text{ means } f = g, \text{ everywhere.}$$

How are these equalities of integral quantities helpful with our physics? I'm going to build a bit more scaffolding, and then revisit this when we talk about Gauss's Law in the section on Electric Fields.

The Directional Derivative and the Gradient

Imagine we are working with a function of x and y only. Suppose we have a vector differential change in position away from a point:

$$\vec{dl} = dx\hat{x} + dy\hat{y}$$

Then we can infer the differential change in f, that is $df = f(x + dx, y + dy) - f(x, y)$ by considering the step in f due to the change in x and the step in f due to the change in y:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

which we can also represent as the following dot product:

$$df = \vec{dl} \cdot \left[\frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y}\right]$$

This is the differential change in f due to an infinitessimal step of length dl in the direction of $\vec{dl}$. But that direction can be specified by the unit vector $\hat{n}$ where $\vec{dl} = dl\hat{n}$. We can rewrite this as the *directional derivative*, the rate of change of f with respect to distance in the direction of $\hat{n}$

$$\frac{df}{dl} = \hat{n} \cdot \left[\frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y}\right]$$

We can rewrite the directional derivative of f as

$$\frac{df}{dl} = \hat{n} \cdot (\vec{\nabla} f) \tag{11}$$

where the gradient operator, or simply the *gradient* is

$$\vec{\nabla} = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y}$$

Through the directional derivative, we can develop an intuitive understanding of the gradient of a function. Consider the function $f = \sqrt{x^2 + y^2}$. It is left as an exercise to show that

$$\vec{\nabla} f = \frac{x\hat{x} + y\hat{y}}{\sqrt{x^2 + y^2}}$$

so the gradient of this particular function is a unit vector that points radially away from the origin at the point where the gradient is taken. In other words, at any location (x, y) (except the origin), $\vec{\nabla} f = \hat{s}$, the radial unit vector in cylindrical coordinates.

In the left panel of the figure below, I show $\vec{\nabla} f$ on three surfaces of constant f. These are circles centered on the origin because the function is just the radial distance from the origin to that point. In other words, $f(x, y) = s$, where s is the radial distance in cylindrical coordinates.

In the middle panel of the figure below, I zero in on one point. At that location, the gradient is a vector pointing radially outwards. The directional derivative, at that point, depends upon the direction of $\hat{n}$ in Equation 11. The fan of vectors are (a subset) of different $\hat{n}$ for which I want to determine the directional derivative of f. The vector in red is one of those vectors, and is also the gradient at that point (because of the way I have set up this situation, the gradient of f is also a unit vector).

In practice, this is a very simple process, though the implementing the math in a computer program is sometimes a bit cumbersome. For this exercise, I took 500 $\hat{n}$ vectors distributed evenly in angle for the 2π degrees between the direction of the gradient, and itself, sixteen of which I show here. The graph in the right hand panel shows the directional derivative as a function of direction with respect to the gradient. The angles corresponding to the gradient and the blue vector are indicated with the red and blue lines on that plot.

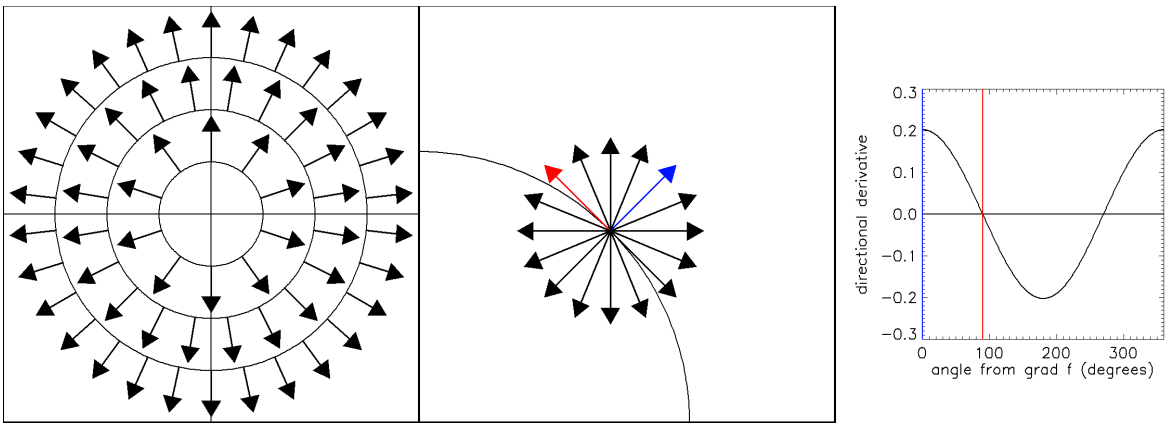


Figure 4: Demonstration of gradient and directional derivative using $f = \sqrt{x^2 + y^2}$.

What do we get from this? The directional derivative is the rate of change of the function in a direction specified by $\hat{n}$ in Equation 11. Its maximum occurs when $\hat{n}$ is parallel to $\vec{\nabla} f$, and its minimum occurs when $\hat{n}$ is anti-parallel to $\vec{\nabla} f$. When $\hat{n}$ is perpendicular to $\vec{\nabla} f$, the directional derivative is zero.

Considering the previous paragraph, it is clear that the gradient of a function is perpendicular to surfaces of constant value of that function. A corollary to that is that infinitessimal steps perpendicular to the gradient means moving along a surface of constant value of that function.

Everything we have done here generalizes to three dimensions. The gradient in a three dimensional Cartesian coordinate system is

$$\vec{\nabla} = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}$$

and the directional derivative is

$$\frac{d}{dl} = \hat{n} \cdot \vec{\nabla}$$

The gradient can be expressed in cylindrical and spherical coordinates. There is a complication that it is important to remember. In Cartesian coordinates $\hat{x}$ is a constant and does not vary with position, and the same is true for $\hat{y}$ and $\hat{z}$, which gives is some freedom in terms of how to write the gradient, namely

$$\hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z} = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}$$

In the case of cylindrical coordinates, for example, both $\hat{s}$ and $\hat{\theta}$ depend on θ, and so

$$\frac{\partial}{\partial \theta}\hat{\theta} \neq \hat{\theta}\frac{\partial}{\partial \theta} \quad \text{and} \quad \frac{\partial}{\partial x}\hat{\theta} \neq \hat{\theta}\frac{\partial}{\partial x}$$

The Divergence and the Divergence Theorem (or Gauss's Theorem)

We define the *flux* of a vector field $\vec{A}$ through a surface S to be

$$\Phi = \int_S \vec{A} \cdot \vec{da} \tag{12}$$

where $\vec{da}$ is $\hat{n}da$, and $\hat{n}$ is perpendicular to the surface and da is a differential area element. Though a scalar, the flux has a kind of directionality in that if I decide that $\hat{n}$ is $\hat{z}$ on a surface in the $z = 0$ plane, and calculate Φ, I get the negative of the flux I get if I pick $\hat{n} = -\hat{z}$. Of course surfaces can be of arbitrary shape, but I picked that simple example to make the point clear.

If the surface is closed, then we can speak of the *outward* flux, where we use the *outward* normal vector, or the *inward* flux, where we use the *inward* normal vector. The outward flux will be the negative of the inward flux.

Let us consider a closed surface which is a *tiny* cube centered on the point (x, y, z), with faces orthogonal to the $\hat{x}$, $\hat{y}$, and $\hat{z}$ directions, and offset from the the point by $h/2$ (e.g., one of the two faces perpendicular to $\hat{x}$ is a square of area h^2 located in the $x - h/2$ plane. Note we will be taking the limit as $h \to= 0$ to be small enough that we can evaluate $\vec{A}$ in the center of each square. The (outward) fluxes through each of the six cube faces are

$$\begin{aligned}
\Delta\Phi_{x-h/2} &= -\hat{x} \cdot \vec{A}(-dx/2, y, z)h^2 \\
\Delta\Phi_{x+h/2} &= \hat{x} \cdot \vec{A}(h/2, y, z)h^2 \\
\Delta\Phi_{y-h/2} &= -\hat{y} \cdot \vec{A}(x, -h/2, z)h^2 \\
\Delta\Phi_{y+h/2} &= \hat{y} \cdot \vec{A}(x, h/2, z)h^2 \\
\Delta\Phi_{z-h/2} &= -\hat{z} \cdot \vec{A}(x, y, -h/2)h^2 \\
\Delta\Phi_{z+h/2} &= \hat{z} \cdot \vec{A}(x, y, h/2)h^2
\end{aligned}$$

The (outward) flux through the surface of the cube is

$$\Delta\Phi = \Delta\Phi_{x=-h/2} + \Delta\Phi_{x=h/2} + \Delta\Phi_{y=-h/2} + \Delta\Phi_{y=h/2} + \Delta\Phi_{z=-h/2} + \Delta\Phi_{z=h/2}$$

or

$$\begin{aligned}
\Delta\Phi &= \Big[A_x(h/2, y, z) - A_x(h/2, y, z) + A_y(x, h/2, z) - A_y(x, -h/2, z) + A_z(x, y, h/2) - A_z(x, y, h/2)\Big]h^2 \\
&= \left[\frac{A_x(h/2, y, z) - A_x(h/2, y, z)}{h} + \frac{A_y(x, h/2, z) - A_y(x, -h/2, z)}{h} + \frac{A_z(x, y, h/2) - A_z(x, y, h/2)}{h}\right]h^3 \\
d\Phi &= \left[\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right]d\tau
\end{aligned}$$

where in the last step I have used the idea that is is an infinitessimal cube, the expression for partial derivatives as a limit, and $d\tau = h^3$. Now, imagine we have a volume V bound by a surface S. We can break this volume up into cubical volume elements like this, and sum the outward fluxes. Every surface that contributes to a $d\Phi$ that is not part of the outer boundary S of V is also one of the surfaces of an adjacent cube. The outward flux through that surface from one cube is the negative of the outward flux through that surface from the second cube. This is true for every surface of every cube within V that is not part of the boundary S of V. The sum of every $d\Phi$ within V is then the outward flux of A through S. Thus

$$\int_V \frac{d\Phi}{d\tau} d\tau = \int_S \vec{A} \cdot \vec{da} \tag{13}$$

However, we have, from above,

$$\frac{d\Phi}{d\tau} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

The quantity on the RHS is

$$\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = \left[\hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}\right] \cdot \vec{A}$$

The operator in parenthesis on the RHS is the gradient, and so we can write

$$\vec{\nabla} \cdot \vec{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

We call this operation, $\vec{\nabla} \cdot \vec{A}$, the *divergence*. We can thus re-express Equation 13 as

$$\int_V \vec{\nabla} \cdot d\tau = \int_S \vec{A} \cdot \vec{da} \tag{14}$$

where the surface integral is understood to be the outward flux (in other words $\vec{da}$ points outwards). This is called by some the Divergence Theorem and by others Gauss's Theorem.

Looked at as a quantity, divergence is a source of flux. If field lines are locally divergent, then that region is a source of flux. Likewise, if they are locally convergent, then that region is a sink for flux.

The Curl and Stokes' Theorem

The vector field in the three panels in the figure below is meant have a zero $z-$component, and no variation of its $x-$ and $y-$components with z. I am showing it on a z =constant plane, with $\hat{z}$ pointing out of the page and . One will often here people describe such a field as having *rotation*.

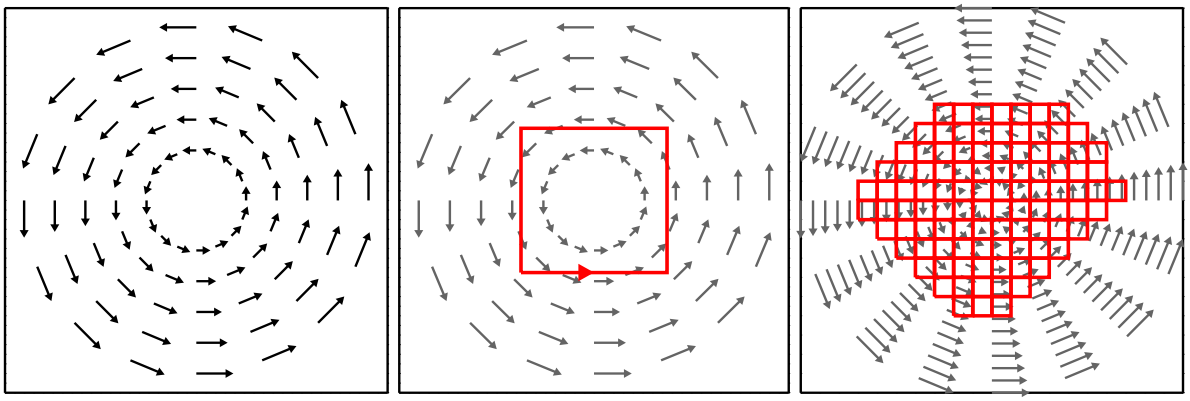


Figure 5: Vector field with rotation.

Ignore the left and middle panels. We are going to focus on the right panel. We can define a quantity, *rotation*, as

$$R = \oint_{\Gamma} \vec{A} \cdot \vec{dl}$$

where this is understood to be the closed path integral of $\vec{A}$ on the contour Γ which bounds a surface S. If I consider one of the little squares in the right hand pane, and determine R with the understanding I am carrying out the path integral in the counter-clockwise direction (meaning the right hand rule would yield $\hat{z}$ for the path integral direction). If I say the square is h by h, and centered on the point (x, y, z), then there are four contributions to the rotation for that little square:

$$\Delta R = h\vec{A}(x + \frac{h}{2}, y, z) \cdot \hat{y} - h\vec{A}(x, y + \frac{h}{2}, z) \cdot \hat{x} - h\vec{A}(x - \frac{h}{2}, y, z) \cdot \hat{y} + h\vec{A}(x, y - \frac{h}{2}, z) \cdot \hat{x}$$

which we can rewrite as

$$\Delta R = hA_y(x + \frac{h}{2}, y, z) - hA_x(x, y + \frac{h}{2}, z) - hA_y(x - \frac{h}{2}, y, z) + hA_x(x, y - \frac{h}{2}, z)$$

or

$$\Delta R = \left[\frac{A_y(x + \frac{h}{2}, y, z) - A_y(x - \frac{h}{2}, y, z)}{h} - \frac{A_x(x, y + \frac{h}{2}, z) + A_x(x, y - \frac{h}{2}, z)}{h} \right] h^2$$

If we imagine the square is differentially small, then we have, the rotation for our little square is

$$\Delta R = \left[\frac{\partial A_y}{\partial x} - \frac{\partial Ax}{\partial y} \right] h^2$$

Therefore,

$$\Delta R = (\vec{\nabla} \times \vec{A})_z da_z$$

If we imagine a closed contour in this z =constant plane, we can fill it with these little squares. Contributions to the sum of the contributions to the rotation from shared internal sides of adjacent little squares cancel (because the direction of path integration is counter-clockwise), and thus

$$\sum \Delta R = \oint \vec{A} \cdot \vec{dl}$$

where the direction of contour integration is counter-clockwise. At the same time

$$\sum \Delta R = (\vec{\nabla} \times \vec{A})_z da_z$$

In he limit of small h, the above two equations mean

$$\oint \vec{A} \cdot \vec{dl} = \int (\vec{\nabla} \times \vec{A})_z da_z$$

It is straightforward, if cumbersome, to show that considering how this works on an arbitrary surface, we have *Stokes' Theorem* which relates the curl of a vector and its path integrals:

$$\int_{surface} \left(\vec{\nabla} \times \vec{A}\right) \cdot \mathrm{d}\vec{a} = \oint_{contour} \vec{A} \cdot \vec{dl}$$

Here, the *contour* is the (obviously closed) boundary of the *surface*, and $\mathrm{d}\vec{l}$ represents an infinitesimal displacement along the contour. For Stokes' Theorem, the surface integral is calculated such that the direction of contour integration and the direction of the vector surface element $\vec{da}$ are correctly related by the right hand rule.

Let the Physics Inform the Math

I'm a fan of a concept I call *let the physics inform the math.* What I mean is for example Gauss's Law is just Gauss's Theorem from math, and it is correct and you can use it as a step in a math derivation. So if I were to write down the volume integral that gives you the electric field of a uniform sphere of charge for a point outside the spherical shell we talked about in class (which is $\vec{E} = Q_{total}\hat{r}/(4\pi\epsilon_\circ r^2)$, I could then quote Gauss's Law (which gave me the above result) and say "*therefore this integral is* $Q_{total}\hat{r}/(4\pi\epsilon_\circ r^2)$*.*" Similarly, if someone were to ask you "*what is the curl of*" $\hat{r}/r^2$? .... you could work it out or you could say "*that's the spatial dependence of a well-known static electric field so the curl of that function is zero... QED.*" Maybe don't take that to the bank with instructors on tests and assignments who often have pedagogical reasons for wanting you to work something out in a specific way, but if you're careful (and correct!!) about how you frame your argument it's a completely valid answer.

The Divergence and Curl in Curvilinear Coordinates

In **cylindrical** coordinates in *very nearly* every instance we will be working with cylindrically symmetric distributions of charge or current, and therefore the potentials and fields will also be cylindrically symmetric. Thus, for *almost all* of this material, when working in cylindrical coordinates, we can use the following (*remember, these are not the complete expressions*):

$$\vec{\nabla} f = \frac{\partial f}{\partial s}\hat{s} + \frac{\partial f}{\partial z}\hat{z} \qquad \vec{\nabla}\cdot\vec{A} = \frac{1}{s}\frac{\partial}{\partial s}\Big(sA_s\Big) + \frac{\partial}{\partial z}A_z$$

$$\vec{\nabla}\times\vec{A} = -\frac{\partial A_\phi}{\partial z}\hat{s} + \left(\frac{\partial A_s}{\partial z} - \frac{\partial A_z}{\partial s}\right)\hat{\phi} + \frac{1}{s}\frac{\partial (sA_\phi)}{\partial s}\hat{z} \qquad \nabla^2 f = \frac{1}{s}\frac{\partial}{\partial s}\left(s\frac{\partial f}{\partial s}\right) + \frac{\partial^2 f}{\partial z^2}$$

Please note that here we are adhering to the formalism adopted by Griffiths, namely in cylindrical coordinates the azimuthal angle is referred to as ϕ and the radial distance from the $z-$axis is referred to as s. As well, in **spherical** coordinates, we use the essentially ubiquitous convection of referring to the polar and azimuthal angles as θ and ϕ, respectively, and the radial distance from the origin as r. Herein, *every* instance we will be working with spherically symmetric distributions of charge or current, and therefore the potentials and fields will also be spherically symmetric. Thus, for *all* of this material, when working in spherical coordinates, we can use the following (*again, remember, these are not the complete expressions*):

$$\vec{\nabla} f = \frac{\partial f}{\partial r}\hat{r} \qquad \vec{\nabla}\cdot\vec{A} = \frac{1}{r^2}\frac{\partial (r^2 A_r)}{\partial r}$$

$$\vec{\nabla}\times\vec{A} = -\frac{1}{r}\frac{\partial (rA_\phi)}{\partial r}\hat{\theta} + \frac{1}{r}\frac{\partial (rA_\theta)}{\partial r}\hat{\phi} \qquad \nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right)$$

The vector and Calculus identities and laws we will be using most frequently include the triple product

$$\vec{A}\times\Big(\vec{B}\times\vec{C}\Big) = \vec{B}\Big(\vec{A}\cdot\vec{C}\Big) - \vec{C}\Big(\vec{A}\cdot\vec{B}\Big)$$

Helmholtz's Theorem

Helmholtz's Theorem addresses a fundamental property of all vector functions, namely that any such function can be constructed as the sum of the gradient of a scalar and the curl of a vector (I cannot find a reasonably penetrable proof of this, but I am working on it and hope to have something pithy on this soon):

$$\vec{A} = \vec{\nabla} F + \vec{\nabla}\times\vec{G} \tag{15}$$

Heaviside Function

In the late 1970s and early 1980s I used to listen to music by a rock band named Oliver Heaviside (in later years they were known as the Partland Brothers). In 1982, they won a big contest held by Q107, one of Toronto's premier rock stations. In my youth, Oliver Heaside was a rock band!

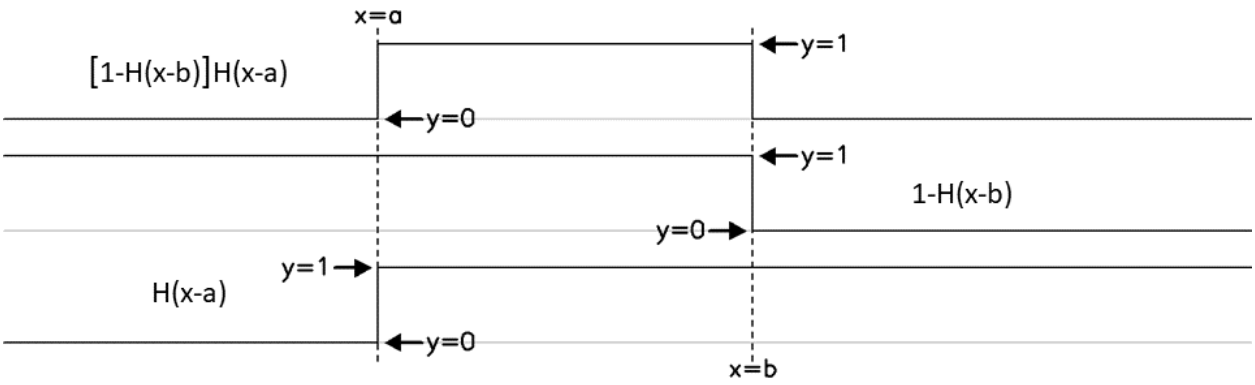
In 1985-1986 I was a fourth year student doing my thesis project in ionospheric physics. I remember reading that the ionosphere was also called the Heaviside Layer. In 1986, I went and saw Cats at Pantages in Toronto: one of the songs was The Journey to the Heaviside Layer. In my *almost* youth, Oliver Heaviside had something to do with the ionosphere (and, I guess, heaven).

It turns out Oliver Heaviside, the band, was named after Oliver Heaviside, the self-educated Scottish physicist and mathematician. And for the cats in Cats, the Heaviside layer was heaven. Heaviside did foundational work in vector calculus, the formulation of Maxwell's Equations, and the existence of an ionized part of Earth's upper atmosphere called the Heaviside layer, and later the ionosphere. While his contributions were monumental, I'm not sure why history chose to immortalize him in this way... our field named a simple but profoundly useful function in his honor: the Heaviside function.

The Heaviside function is defined as follows:

$$H(u) = 1 \text{ for } u > 0 \quad \text{and} \quad H(u) = 0 \text{ for } u < 0$$

In the following plot, I show $H(x-a)$, $1-H(x-b)$, and $H(x-a)[1-H(x-b)]$.



This function has some interesting properties. For example, if we want to represent a piecewise continuous e.g. charge distribution, such as $\rho(x) = 0$ for $x < a$, $\rho(x) = \rho_\circ$ for $a < x < b$, and $\rho(x) = 0$ for $x > b$, we can simply say $\rho(x) = H(x-a)[1-H(x-b)]\rho_\circ$.

Dirac Delta Function

The Dirac delta function is an example of a broad class of generalized functions that are defined by what they do as part of the integrand of a definite integral. Without addressing the extensive mathematical intricacies around the concept, two two functions $\Phi(x)$ and $\Psi(x)$ are said to be *equal in the sense of generalized functions* if

$$\int_{-\infty}^{\infty} \Phi(x) f(x) dx = \int_{-\infty}^{\infty} \Psi(x) g(x) dx$$

for all reasonably behaved functions $f(x)$. Such functions are defined *under an integral*, and for any such function, there is always an infinite number of functions that satisfy the equality. An example which is straightforward to understand and conceptualize is

$$\lim_{k\to\infty} \sin(x) = 0 \text{ in the sense of generalized functions}$$

In physics, probably the most widely known and utilized generalized function is the Dirac delta function, represented symbolically by $\delta(x)$. It is defined according to the following

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0)$$

As I said, there are an infinite number of *generalized* functions each of which *are* the Dirac delta function. For example, consider the function

$$\Psi(x) = \lim_{a\to 0} \Psi(x,a) \quad \text{where} \quad \Psi(x,a) = \frac{H(x+a) - H(x-a)}{2a}$$

where $H(x)$ is the Heaviside function, and the left panel in Figure is a plot of $\Psi(x)$. We are free to reverse the order of taking the limit and integrating, so

$$\int_{-\infty}^{\infty} \Psi(x) f(x) dx = \lim_{a\to 0} \int_{-\infty}^{\infty} \Psi(x,a) f(x) dx = \lim_{a\to 0} \frac{1}{2a} \int_{-a}^{a} f(x) dx$$

Based on the idea of integrals as averages, we can say, first

$$\int_{-a}^{a} f(x) dx = 2a f^*$$

where f^* is the average of f on the interval (-a,a). Therefore

$$\int_{-\infty}^{\infty} \Psi(x) f(x) dx = \lim_{a\to 0} \frac{1}{2a} 2a f^* = \lim_{a\to 0} f^* = f(0)$$

Therefore

$$\int_{-\infty}^{\infty} \Psi(x) f(x) dx = f(0)$$

and so, in the sense of generalized functions,

$$\Psi(x) = \delta(x)$$

Dummy Variables

You may have noticed as I have worked through these concepts I would sometimes use primed (e.g., x', $\vec{r}'$, dx', da', $d\tau'$) and sometimes unprimed (e.g., x, $\vec{r}$, dx, da, $d\tau$). While it may seem a bit arbitrary, my choices are in general quite purposeful.

The concept of the *dummy variable* is straightforward. Under an integral, the dummy variable(s) is(are) integrated *away*. So if we look at the following

$$F(x) = \int_{x_0}^{x_1} xx'dx'$$

we integrate over x'. The result is

$$F(x) = x\frac{x'^2}{2}\Big|_{x'=x_0}^{x'=x_1} = \frac{x(x_1^2 - x_0^2)}{2}$$

It is important to be able to look at integrals and determine what they are a function of. This becomes critically important when we deal with Coulomb's Law and the Biot-Savart Law.

Order of Operations

We can generally count on being able to reverse the order of operations, so that we have

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y}f(x,y) = \frac{\partial}{\partial y}\frac{\partial}{\partial x}f(x,y),$$

$$\vec{\nabla}\cdot\int_V \vec{A}(\vec{r},\vec{r}')d\tau' = \int_V \vec{\nabla}\cdot\vec{A}(\vec{r},\vec{r}')d\tau'$$

and

$$\int_V \lim_{h\to 0} f(h,\vec{r}')d\tau' = \lim_{h\to 0}\int_V f(h,\vec{r}')d\tau'$$

For pedagogical reasons, we do not work out many integrals, nor for that matter do many questions that have numerical answers (except for the many questions for which the answer is zero). That said, because our goal here is to deepen our conceptual understanding of classical physics, including electromagnetism, we will be working a lot with integral expressions. In many cases these integral quantities are *definite* and are themselves functions, such as in the cases of the (Coulomb) electric and (Netwon) gravitational field:

$$\frac{\vec{F}(\vec{r})}{q} = \frac{1}{4\pi\epsilon_\circ}\int \rho_{charge}(\vec{r}')\frac{(\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3}d\tau' \quad \text{and} \quad \frac{\vec{F}(\vec{r})}{m} = G\int \rho_{mass}(\vec{r}')\frac{(\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3}d\tau'$$

In each case the field is a vector function of position, and the integral over the (static) charge distribution is done using the *primed* variable of position. When you get or set up an integral it is usually a good investment to take the time necessary to fully understand which is(are) the *dummy* variable(s) over which the integration is performed. The integral is not a function of the dummy variables, but often is a function of one or more other variables.

Vector and Vector Calculus Identities

Note: We will be using some (but not a lot of) vector identities and vector calculus relations and theorems. In **Cartesian** coordinates we have

$$\vec{dl} = dx\hat{x} + dy\hat{y} + dz\hat{z}$$

$$\vec{\nabla}\cdot\vec{A} = \frac{\partial}{\partial x}A_x + \frac{\partial}{\partial y}A_y + \frac{\partial}{\partial z}A_z \qquad \vec{\nabla}\times\vec{A} = \left(\frac{\partial}{\partial_y}A_z - \frac{\partial}{\partial_z}A_y, \frac{\partial}{\partial_z}A_x - \frac{\partial}{\partial_x}A_z, \frac{\partial}{\partial_x}A_y - \frac{\partial}{\partial_y}A_x\right)$$

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

The following are fundamentally important, and routinely used, and should on one hand be committed to memory, and on the other be derivable by you:

$$\vec{\nabla}\times\left(\vec{\nabla}f\right) = 0 \quad \text{and} \quad \vec{\nabla}\cdot\left(\vec{\nabla}\times\vec{A}\right) = 0$$

There are a number of product rules for vector calculus, including

$$\vec{\nabla}fg = f\vec{\nabla}g + g\vec{\nabla}f \quad \text{and} \quad \vec{\nabla}\cdot f\vec{A} + f\vec{\nabla}\cdot\vec{A} + \vec{A}\cdot\vec{\nabla}f$$

both of which are follow straightforwardly from the product rule for scalar functions. The following

$$\vec{\nabla}\times\left(f\vec{A}\right) = f\vec{\nabla}\times\vec{A} - \vec{A}\times\vec{\nabla}f$$

is slightly more difficult to derive (see questions), but proves equally valuable.

We will be using some (but not a lot of) vector identities and vector calculus relations and theorems. In **Cartesian** coordinates we have

$$\vec{dl} = dx\hat{x} + dy\hat{y} + dz\hat{z}$$

$$\vec{\nabla}\cdot\vec{A} = \frac{\partial}{\partial x}A_x + \frac{\partial}{\partial y}A_y + \frac{\partial}{\partial z}A_z \qquad \vec{\nabla}\times\vec{A} = \left(\frac{\partial}{\partial_y}A_z - \frac{\partial}{\partial_z}A_y, \frac{\partial}{\partial_z}A_x - \frac{\partial}{\partial_x}A_z, \frac{\partial}{\partial_x}A_y - \frac{\partial}{\partial_y}A_x\right)$$

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

The vector and Calculus identities and laws we will be using most frequently include the triple product

$$\vec{A}\times\left(\vec{B}\times\vec{C}\right) = \vec{B}\left(\vec{A}\cdot\vec{C}\right) - \vec{C}\left(\vec{A}\cdot\vec{B}\right)$$

and its very important (for our purposes) vector Calculus equivalent

$$\vec{\nabla}\times\left(\vec{\nabla}\times\vec{A}\right) = \vec{\nabla}\left(\vec{\nabla}\cdot\vec{A}\right) - \nabla^2\vec{A}$$

The following are fundamentally important, and routinely used, and should on one hand be committed to memory, and on the other be derivable by you:

$$\vec{\nabla} \times \left(\vec{\nabla} f \right) = 0 \quad \text{and} \quad \vec{\nabla} \cdot \left(\vec{\nabla} \times \vec{A} \right) = 0$$

There are a number of product rules for vector calculus, including

$$\vec{\nabla} fg = f\vec{\nabla} g + g\vec{\nabla} f \quad \text{and} \quad \vec{\nabla} \cdot f\vec{A} + f\vec{\nabla} \cdot \vec{A} + \vec{A} \cdot \vec{\nabla} f$$

both of which follow straightforwardly from the product rule for scalar functions. The following

$$\begin{aligned}
\vec{\nabla} \times \left(f\vec{A} \right)_x &= \frac{\partial}{\partial y} f A_z - \frac{\partial}{\partial} f A_y \\
&= \left(f \frac{\partial}{\partial y} A_z - f \frac{\partial}{\partial z} A_y \right) + A_z \frac{\partial f}{\partial y} - A_y \frac{\partial f}{\partial z} \\
&= f \left(\vec{\nabla} \times \vec{A} \right)_x + A_z \left(\vec{\nabla} f \right)_y - A_y \left(\vec{\nabla} f \right)_z \\
&= f \left(\vec{\nabla} \times \vec{A} \right)_x - \left(\vec{A} \times \vec{\nabla} f \right)_x
\end{aligned}$$

Doing the same for the $y-$ and $z-$components of $\vec{\nabla} \times f\vec{A}$ gives the following product rule:

$$\vec{\nabla} \times \left(f\vec{A} \right) = f\vec{\nabla} \times \vec{A} - \vec{A} \times \vec{\nabla} f$$

The Math Underpinning the Scalar Potential

The idea that if a vector field is curl free it can be represented as the gradient of a scalar field is widely used in E&M. We can show that this is true for a curl free field by constructing a function whose gradient is that curl free field. This is just a multi-dimensional extension of the Second Fundamental Theorem of Calculus. Consider in one dimension, the following function

$$F(x) = \int_{x_\circ}^{x} f(x')dx'$$

How do you prove that $f(x)$ is the derivative of $F(x)$? One way is as follows. Consider

$$\frac{d}{dx}F(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \lim_{h\to 0}\frac{\int {x_\circ}^{x+h} f(x')dx' - \int {x_\circ}^{x} f(x')dx'}{h} = \lim_{h\to 0}\frac{\int_x^{x+h} f(x')dx'}{h}$$

Then, by the Mean Value Theorem, and letting $\tilde{x}$ be either *the* or *one of the* x-vales between x and $x+h$ where $f(x)$ assumes its mean value on the interval (x,$x+h$), and using the definition of average (or integration), then

$$\frac{d}{dx}F(x) = \lim_{h\to 0}\frac{\int_x^{x+h} f(x')dx'}{h} = \lim_{h\to 0}\frac{hf(\tilde{x})}{h} = f(x)$$

If $\nabla \times \vec{A} = 0$, then the integral

$$\Psi(x,y,z) = \int_{(x_\circ,y_\circ,z_\circ)}^{(x,y,z)} \vec{A}\cdot d\vec{l}$$

is a function (Do you know why?). More than that, it's a function whose gradient is $\vec{A}$. To show this, you just need to show that the three relevant partial derivatives of Ψ are

$$\frac{\partial}{\partial x}\Psi = A_x, \quad \frac{\partial}{\partial y}\Psi = A_y \text{ and } \frac{\partial}{\partial z}\Psi = A_z$$

This involves doing for each of x, y, and z the same thing as we did in the one dimensional equivalent. For example, for the partial derivative with respect to x,

$$\frac{\partial}{\partial x}\Psi = \lim_{h\to 0}\frac{\int_{(x,y,z)}^{(x+h,y,z)} \vec{A}\cdot d\vec{l}}{h}$$

This contour integral is path independent, so we can pick a path we find convenient. One logical choice is to follow the contour that goes from (x,y,z) to (x+h,y,z) along the straight line joining the two points. This is displacement in the x-direction only, so the integral reduces to the length of the line segment (h) times the average of the integrand on that segment ($A_x(\tilde{x},y,z)$) where $\tilde{x}$ has the same meaning as it did in the one dimensional case. Taking the limit, we have

$$\frac{\partial}{\partial x}\Psi = Ax$$

Nothing about this was specific to x, so the same relationship holds for the y and z derivatives...

$$\frac{\partial}{\partial x}\Psi = A_x, \quad \frac{\partial}{\partial y}\Psi = A_y \text{ and } \frac{\partial}{\partial z}\Psi = A_z$$

Not only have we shown that if $\nabla \times \vec{A} = 0$ then $\vec{A}$ is the gradient of a scalar function, we also know how to construct a function whose gradient is $\vec{A}$. The electric and gravitational potential that you are used to are specific examples of this.

Some Thoughts on Tensors

For this level of E&M we have to use vector identities. There are a few things I want you to know and be able to show on, e.g., a test. Examples include the curl of a gradient is zero which you can *prove* to my satisfaction by saying something like the following: "The curl of a gradient is zero because its components are of the form

$$[\nabla \times (\nabla A)]_x = \left[\frac{\partial^2}{\partial y \partial z} - \frac{\partial^2}{\partial z \partial y}\right] A$$

and we can switch the order of first partial derivatives."

I will provide you with the more complicated identities on the formula sheets for each test and the exam, so you do not *need* to be able to derive them on tests. Still, there are ways of dealing with these identities that are fairly easy to remember and that you can use to derive them. *This is likely not required for whichever course you are in*, but I would recommend you take a bit of time to think about this. It will help you with thinking about some necessary electromagnetics concepts, including the Maxwell Stress Tensor. While the *entities* I deal with here are tensors, the ways I use them are simply enabled by properties of matrices, and so I do not want to address the issue of what is and is not a tensor, types of tensors, etc. For our purposes, these are meant to make certain vector operations more tractable.

To begin, we know that vectors can be thought of as matrices. In fact, the vector (A_x, A_y, A_z) is really just a collection of three numbers that we use in certain proscribed ways (the rules and properties of vectors described above). We can represent this by stating a length and a direction (e.g., "such and such long and pointing in such and such direction"), by specifying the three components, or more abstractly by the symbolic representation $\vec{A}$ by which we mean something that has a length and a direction. We can also represent this vector, perfectly generally, as A_i, on the understanding that i has 3 values (1, 2, and 3), and that $A_1 = A_x$, $A_2 = A_3$, and $A_3 = A_z$.

I can generalize this to multiple dimensions, using, e.g., A_{ij} and A_{ijk} to two and three dimensional represent collections of numbers where, again, 1, 2, and 3, correspond to [the] x, y and z [directions].

An example you are undoubtedly familiar with is the Kronecker Delta symbol, δ_{ij}, which for our purposes is 3X3 matrix where $\delta_{ij} = 0$ where $i \neq j$ and 1 where $i = j$. It is not the custom to think of this as the identity matrix, but that is what it is. The statement $I = \delta_{ij}$ (where I is the identity matrix) is a reasonable thing to write down, though the representation conventions are different. The quantity on the left hand side is the matrix and so too is the quantity on the right hand side. When I write down δ_{ij}, I mean all nine numbers. This is a collection of numbers, not the ijth component of the matrix.

Further, we can add a convention that we can multiply these quantities and that *repeated indices represent summation*. This is often called the Einstein Summation Convention (I do not know why). So the dot product, in this convention, is

$$\vec{A} \cdot \vec{B} = A_i B_i$$

Getting used to this is easier if you focus on rigidly applying the convention and making sure of the logical equivalence between both sides of the equation. On the right hand side, we have $A_1B_1 + A_2B_2 + A_3B_3$ which is the dot product and thus equivalent to what is on the left hand side. Similarly

$$\delta_{ii} = 3$$

The left hand side is $\delta_{11} + \delta_{22} + \delta_{33}$ or $1 + 1 + 1$, and thus the equivalent of the right hand side. Moving to

three dimensions, the Levi-Cevita symbol, ϵ_{ijk}, is defined such that ϵ_{ijk}=0 where $i = j$ or $i = k$ or $j = k$, ϵ_{ijk}=1 if $ijk = 123$ or its two circular permutations ($ijk = 312$ or $ijk = 231$), and ϵ_{ijk}=-1 if $ijk = 213$ or its two circular permutations ($ijk = 321$ or $ijk = 132$). That seems like a lot, but it is really just the following. ϵ_{ijk}=0 if any two of the indices (or more) have the same value, ϵ_{ijk}=1 if ijk is a 123 or a circular permutation of 123, and otherwise ϵ_{ijk}=-1.