PHYS2796 Introduction to Modern Physics (Spring 2015)

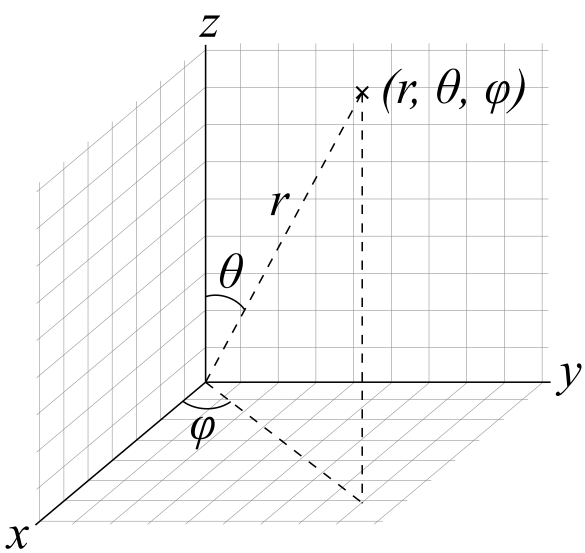
Notes on Mathematics Prerequisites

Jim Napolitano, Department of Physics, Temple University
January 7, 2015

This is a brief summary of material on mathematics with which students should be familiar for Physics course 2796 *Introduction to Modern Physics* in Spring 2015.

Coordinate systems and vectors in three spatial dimensions

We locate a point in space according to either its Cartesian coordinates (x, y, z) or its spherical coordinates (r, θ, ϕ) as defined in the usual (by physicists) manner as shown here:



$$x = r \sin\theta \cos\phi \tag{1a}$$
$$y = r \sin\theta \sin\phi \tag{1b}$$
$$z = r \cos\theta \tag{1c}$$

$$r^2 = x^2 + y^2 + z^2 \tag{2a}$$
$$\tan\theta = \frac{\sqrt{x^2 + y^2}}{z} \tag{2b}$$
$$\tan\phi = \frac{y}{x} \tag{2c}$$

A shorthand notation $\vec{\mathbf{r}}$ is used to denote the position of a particle. That is, $\vec{\mathbf{r}}$ can denote either (x, y, z) or (r, θ, ϕ). This is a special case of a *vector* in three dimensions. A vector is essentially a quantity that has both direction and magnitude (or "length.")

We generalize the notion of a vector by defining the *unit vector*, which has unit magnitude (that is, length equal to the dimensionless value 1) but points in the direction in which a particular coordinate increases. We put little "hats" over the symbol to denote a unit vector. For example, $\hat{\mathbf{x}}$ is the vector with length$= 1$ that points in the direction of the positive x-axis, and $\hat{\mathbf{r}}$ is the unit vector that points radially out from the origin. We therefore have

$$\vec{\mathbf{r}} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}} \tag{3a}$$
$$= r\hat{\mathbf{r}} \tag{3b}$$

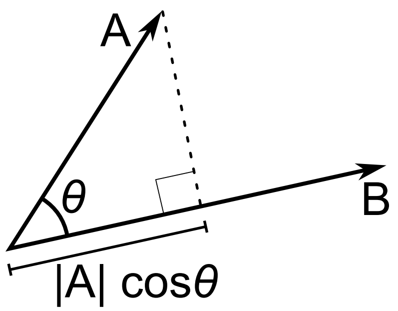
It is very important to realize that unlike the Cartesian unit vectors, the spherical unit vectors point in different directions at different points in space.

In general, a vector $\vec{\mathbf{A}}$ refers to three quantities (A_x, A_y, A_z) and we write

$$\vec{\mathbf{A}} = A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}} \tag{4}$$

where (A_x, A_y, A_z) have their spherical counterparts (A_r, A_θ, A_ϕ) that follow the same relationships as in Equations 1 and 2. In fancy terminology, this means that $\vec{\mathbf{A}}$ "transforms like a vector" when it comes to the rotational transformation in three dimensions.

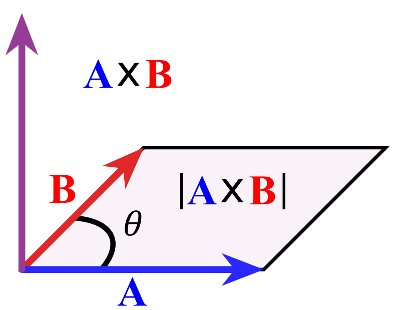
We can form the *inner product* or *scalar product* or *dot product* between two vectors $\vec{\mathbf{A}}$ and $\vec{\mathbf{B}}$, written as $\vec{\mathbf{A}} \cdot \vec{\mathbf{B}}$, using a simple geometrical definition. That is, $\vec{\mathbf{A}} \cdot \vec{\mathbf{B}}$ is just the number one gets by multiplying the length of $\vec{\mathbf{B}}$ by the length of the projection of $\vec{\mathbf{A}}$ onto $\vec{\mathbf{B}}$. Trigonometrically, if θ is the angle between the two vectors, and writing $|\vec{\mathbf{A}}|$ for the length of a vector, we have



$$\begin{aligned} \vec{\mathbf{A}} \cdot \vec{\mathbf{B}} &= |\vec{\mathbf{A}}||\vec{\mathbf{B}}| \cos\theta & \text{(5a)} \\ &= A_xB_x + A_yB_y + A_zB_z & \text{(5b)} \end{aligned}$$

where the second relationship follows from the definition of the unit vector and the fact that the unit vectors are all perpendicular (i.e. "orthogonal") to each other. Note that we can always relate the length of a vector to the dot product using $|\vec{\mathbf{A}}|^2 = \vec{\mathbf{A}} \cdot \vec{\mathbf{A}}$.

Another useful definition is the *cross product* of two vectors $\vec{\mathbf{A}}$ and $\vec{\mathbf{B}}$:



$$\begin{aligned} \vec{\mathbf{A}} \times \vec{\mathbf{B}} &= (A_zB_y - A_yB_z)\hat{\mathbf{x}} \\ &+ (A_xB_z - A_zB_x)\hat{\mathbf{y}} \\ &+ (A_yB_x - A_xB_y)\hat{\mathbf{z}} \end{aligned} \tag{6}$$

In this case, the geometric interpretation is a vector that is perpendicular to both $\vec{\mathbf{A}}$ and $\vec{\mathbf{B}}$, in a direction determined by convention according to the "right hand rule", and with a magnitude equal to the area of the parallelogram determined by $\vec{\mathbf{A}}$ and $\vec{\mathbf{B}}$.

Sometimes we will refer to a vector as a *three-vector* to distinguish it from the quantity called a *four-vector* that we will encounter in Special Relativity. Indeed, the notion of a "vector" is a very general one that is used in many formulations of physical theories.

Elements of differential and integral calculus, and series expansions

Above all, you should think of differentiation and integration *physically*. That is, derivatives tell us how some quantity *changes* in some spatial direction or over time, and this change can differ at different points in space or at different times. An integral is a *sum* of a very large (infinite) number of tiny (infinitesimal) pieces; it's no accident that $\int$ looks like an S.

Of course, along the way, you'll learn how to derive (and then can't help but memorize) various calculus formulas. For $y(x) = Ax^\alpha$, we all know that $dy/dx = \alpha A x^{\alpha-1}$ for any value of α. Integration is just the reverse, but with some special cases, for example $\int_1^x (1/t)dt = \ln x$. You can look up derivatives and integrals online or in books, or just use a program like MATHEMATICA to evaluate them.

Understanding the notation is important. For integrals, we will almost never use the "indefinite" integral, also known as the "anti-derivative". Only definite integrals can be interpreted as a sum, so that's what we'll stick with. For derivatives, with a function, say, $y = f(x)$, we might write the derivative as dy/dx or $f'(x)$. If we want to evaluate the derivative at a certain point $x = x_0$, we would write either $dy/dx|_{x=x_0}$ or $f'(x_0)$.

We can literally think of dy/dx as the ratio of two tiny quantities dy and dx. This is an easy way to derive the "chain rule" from your calculus class. If $y = f(u)$ and $u = g(x)$, then

$$\frac{dy}{dx} = \frac{dy}{dx} \times 1 = \frac{dy}{dx}\frac{du}{du} = \frac{dy}{du}\frac{du}{dx} = f'(u)g'(x) \tag{7}$$

Physics often deals with quantities that change with time t, and we have an old notation for derivatives with respect to time. For example, if x is a function of time, i.e. $x = f(t)$, then we write $\dot{x} = f'(t)$. Two dots means two time derivatives, and so on. In other words, if x measures position, then $\dot{x} = v$ is the velocity and $\ddot{x} = \dot{v} = a$ is the acceleration.

Remember that integrals are sums, and we can add up anything we want. So, for example, a line integral $\int_a^b \vec{\mathbf{A}} \cdot d\vec{\mathbf{l}}$ means to add up all the little pieces of the dot product of $\vec{\mathbf{A}}$ with a directional line segment, all along a line from point a to point b. A surface integral $\int_S \vec{\mathbf{A}} \cdot d\vec{\mathbf{S}}$ similarly means to add up all of the little pieces of $\vec{\mathbf{A}} \cdot d\vec{\mathbf{S}}$ over the surface S. Sometimes the line forms a closed loop C, or the surface forms a closed region of volume; in such cases we write $\oint_C \vec{\mathbf{A}} \cdot d\vec{\mathbf{l}}$ or $\oint_S \vec{\mathbf{A}} \cdot d\vec{\mathbf{S}}$.

The "Taylor Expansion" is very important for physical thinking and for theoretical concepts:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2!} + f''(x_0)(x - x_0)^2 + \frac{1}{3!} + f'''(x_0)(x - x_0)^3 + \cdots \tag{8}$$

For example, physically, we very often make use of $(1 + x)^\alpha \approx 1 + \alpha x$ for $x \ll 1$. You should also be familiar with the Taylor expansions (about $x = 0$) for $\cos(x)$, $\sin(x)$, and $\exp(x) \equiv e^x$. These expansions, in fact can be used to derive Euler's formula, namely

$$e^{ix} = \cos(x) + i\sin(x) \tag{9}$$

where $i^2 = -1$. It would be smart for you to derive all of these formulas on your own.

Ordinary differential equations

You *do not* need to have taken a course in differential equations in order to master the material in *Introduction to Modern Physics*. You do, however, need to know what they are, and what it means to have a solution for them. This is not hard. In many cases, it is also easy to guess the solution to a differential equation. You'll learn all about the many different ways to solve these equations, though, when you take the appropriate math course.

A differential equation is just an equation that includes derivatives of some function. A "first order" differential equation involves only first derivatives, whereas a "second" (or higher) order equation involves derivatives up to that order. To "solve" a differential equation means to find a function that satisfies the equation. One needs to specify n "boundary" or "initial" values in order to find the complete solution to an n-th order differential equation.

Here is a simple example. An object of mass m falls, starting at rest, from a height h. What is the function $y(t)$ that measures its height as a function of time? Newton's Second Law "$F = ma$" becomes the differential equation $-mg = m\ddot{y}$. The general solution is simple, just integrate twice to find $y(t) = a + bt - gt^2/2$ for any constants a and b. However $y(0) = h$ means $a = h$ and $\dot{y}(0) = 0$ means that $b = 0$. Therefore $y(t) = h - gt^2/2$.

On your own, try something harder. Add a drag force $-Cv$ and find the velocity $v(t)$.

Remember that you can always use whatever you can think of to guess the solution to an equation. (There are theorems having to do with "uniqueness" but we won't bother with that here.) Suppose, for example, you are faced with the differential equation

$$\frac{d^2\psi}{dx^2} = -k^2\psi(x) \qquad \text{where} \qquad \psi(0) = \psi(L) = 0 \tag{10}$$

The function $\psi(x)$ has the property that taking the derivative twice gives the same function back, but multiplied by $-k^2$. You know *two* such functions, namely $\cos(kx)$ and $\sin(kx)$. You can be formal and write $\psi(x) = A\cos(kx) + B\sin(kx)$, but you quickly realize that the boundary condition implies that $A = 0$, and that $kL = n\pi$ where n is an integer. We'll see this equation often in this course.

That last example made use of the concept of "linearity". A differential equation is called linear if the function and its derivatives only appear to first order. For such equations, it is easy to show that if $f(x)$ and $g(x)$ are both solutions, then any function of the form $Af(x) + Bg(x)$ where A and B are constants is also a solution.

By the way, we are only discussing here differential equations where the function for which we want to solve, depends only on one independent variable, which we've called t or x. These are called "ordinary" differential equations. Otherwise, we're talking about so-called "partial differential equations."

Partial derivatives and partial differential equations

Below is a photograph of a (red) tail attached to a (blue) kite. Forget that the tail is thick, and let's consider how to describe the shape of the tail, as if it were infinitely thin:

If "up" is y, x is to the right, and time is t, then we might write $y = f(x,t)$ to describe the shape of the tail at some arbitrary time. It is a "curve", y as a function of x, that changes with time.

We would write the slope of the curved shape of the tail as $\partial f/\partial x$, that is, the derivative of y with respect to x at some fixed time. Alternatively, we can pick one point located at some x, and ask what is its up or down velocity at any time. This we would write as $\partial f/\partial t$. The funny looking "∂" just means to take the derivative of one variable while keeping the other variable fixed, that is, treated as if it were a constant.

These are called *partial derivatives* and we work with the same way as we work with ordinary derivatives. Just hold all other variables constant when you take the derivative with respect to the one that you care about. Of course, we can form *partial differential equations* out of these, but they can be complicated to solve. We nevertheless will encounter these kinds of equations in this course, but you'll be guided through their solutions.

Another place that partial derivatives show us, is when we consider a function of three dimensional space, for example $\psi(x,y,z)$, which might also be written as $\psi(\vec{\mathbf{r}})$. Consider

$$\boldsymbol{\nabla}\psi = \hat{\mathbf{x}}\frac{\partial\psi}{\partial x} + \hat{\mathbf{y}}\frac{\partial\psi}{\partial y} + \hat{\mathbf{z}}\frac{\partial\psi}{\partial z} \qquad \text{that is,} \qquad \boldsymbol{\nabla} \equiv \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z} \tag{11}$$

The differential operator $\boldsymbol{\nabla}$ is called the *gradient operator* and clearly is some measure of how the function $\psi(x,y,z)$ is changing in space. One can also form the operator

$$\boldsymbol{\nabla}^2\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} \qquad \text{that is,} \qquad \boldsymbol{\nabla}^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{12}$$

which is called the *Laplacian operator*. The very neat form of $\boldsymbol{\nabla}^2$ in Cartesian coordinates, comes about because the unit vectors are constant in space. If we work in spherical coordinates, which in fact we will do in this course, we need to take into account how the unit vectors themselves change when we take the gradient dotted into the gradient. One finds

$$\boldsymbol{\nabla}^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{13}$$

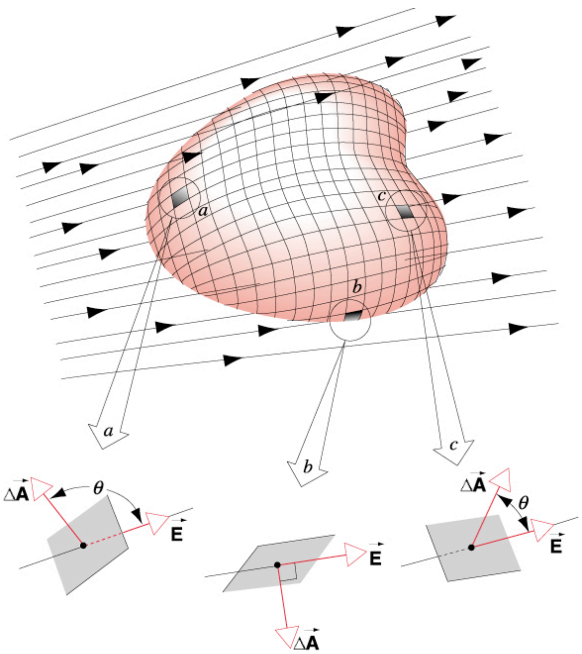
We'll arrive at this point when we try solving the hydrogen atom in wave mechanics, but you'll be taken through the mathematics of that partial differential equation step by step. It's a lot less complicated than it seems.

There is another differential operator, called the *curl* of a *vector field*, which is defined as $\boldsymbol{\nabla}\times\vec{\mathbf{A}}(\vec{\mathbf{r}})$ for some vector function $\vec{\mathbf{A}}(\vec{\mathbf{r}})$, but we won't run into it in this course.

Vector integral calculus and the surface theorems

This is material you'll not likely have seen yet. It is also material that we will not use heavily in this course, but you will certainly see it if you move into more theoretical physics courses.

I'll illustrate the math with a physical example, which in fact applies to very many different physical situations. Consider a closed region of space $\mathcal{R}$ containing some water. You can think of that space enclosed by a porous mesh surface that we'll call $\mathcal{S}$, submerged into a large pool. Water can through the mesh, both into and out of the region $\mathcal{R}$.



We have $\vec{\mathbf{E}}(\vec{\mathbf{r}},t) = \rho(\vec{\mathbf{r}},t)\vec{\mathbf{v}}(\vec{\mathbf{r}},t)$ where ρ is the density and $\vec{\mathbf{v}}$ is the velocity of a small bit of water located at $\vec{\mathbf{r}}$ at a time t. As shown in the figure, the mass of the water that flows through a small surface area $d\vec{\mathbf{A}}$ on the closed surface $\mathcal{S}$ is just $\vec{\mathbf{E}}\cdot d\vec{\mathbf{A}}$. (The direction of $d\vec{\mathbf{A}}$ is chosen so that it always points outwardly perpendicular from $\mathcal{S}$.)

We can now use math to write that *water can neither be created or destroyed*:

$$\frac{d}{dt}\int_{\mathcal{R}} \rho dV = -\oint_{S} \vec{\mathbf{E}}\cdot d\vec{\mathbf{A}} \tag{14}$$

In other words, the only way that we can change the amount of water inside $\mathcal{R}$ is for water to flow into or out of that region through the surface $\mathcal{S}$. The minus sign on the right is because, since $d\vec{\mathbf{A}}$ points outward, a positive integral means that water is leaving $\mathcal{R}$.

Now we can use a "surface theorem", known as Gauss' Theorem or the Divergence Theorem:

$$\oint_{S} \vec{\mathbf{E}}\cdot d\vec{\mathbf{A}} = \int_{\mathcal{R}} \boldsymbol{\nabla}\cdot\vec{\mathbf{E}}\; dV \qquad \text{where} \qquad \boldsymbol{\nabla}\cdot\mathbf{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \tag{15}$$

which means we can write the "law" of conservation of water very succinctly as

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla}\cdot(\rho\vec{\mathbf{v}}) = 0 \tag{16}$$

I've skipped a lot of steps here, not to mention disregarding a proof of the Divergence Theorem, but all I really want to do is get you thinking about the direction one can go with this mathematics. In Quantum Mechanics, this is how you show that Schrödinger's equation conserves probability, how Maxwell's Equations conserve electric charge, and how Einstein's theory of General Relativity conserves momentum and energy.