Vector Calculus Review

Mathew Calkins

December 19, 2025

Abstract

Notes on vector algebra and vector calculus, focused on the needs of classical mechanics and electromagnetism in Galilean (*i.e.*, not relativistic) notation.

Contents

1 Vector Algebra

Vectors are perhaps the must ubiquitous mathematical objects in physics after basic characters like sets and functions. Depending on the context, we could give vectors any of the following increasingly abstract definitions, not all of which are equivalent and not all of which are actually carefully well-defined:

1. Something with magnitude and direction.

2. A triple of real numbers (x, y, z).

3. A tuple of real numbers $(x_1, \ldots, x_n)$.

4. An element of a tangent space to our favorite manifold.

5. Something that transforms like a vector.

6. Something that transforms in the fundamental representation of our favorite group.

7. A ket.

8. An element of a vector space.

Undergraduate physicists famously struggle to reconcile these inequivalent definitions. By the end of this course we will mostly work with Definition 4 per the demands of calculus on manifolds, and as you mature you will eventually find Definition 8 most useful and think of the rest as special cases. At present we are interested in vectors as they appear in undergraduate courses on classical mechanics and electrodynamics, so we will use Definition 3, and taking Definition 2 as a useful special case.

Said plainly, we are interested in n-dimensional Euclidean space $\mathbb{R}^n$. As a set, this is the collection of all ordered tuples $\vec{x} = (x_1, \ldots, x_n)$ where each x_i is a real number. Borrowing a bit of useful notation from pure mathematics, we sometimes write

$$\vec{x} \in \mathbb{R}^n \text{ and } x_i \in \mathbb{R}.$$

Here "$x_i \in \mathbb{R}$" is pronounced "ex eye in are" and means x_i is an element of $\mathbb{R}$, the set of all real numbers. We call x_i the ith component of $\vec{x}$. When it is convenient, we sometimes enumerate the components not as a horizontally typeset list inside of parentheses but as a column vector like

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

1.1 Linear Combination

The most basic thing we can do with vectors is to take linear combinations. Given vectors $\vec{x}, \vec{y}$ and real numbers a, b, we define their linear combination component-wise as

$$a\vec{x} + b\vec{y} := \begin{bmatrix} ax_1 + by_1 \\ \vdots \\ ax_n + by_n \end{bmatrix}.$$

We will sometimes write $:=$ when an equality holds by definition, with $:$ on the side of the new object being defined.

This has a nice geometric interpretation. If we picture the vectors $\vec{x}, \vec{y}$ as arrows, then scaling them individually changes the lengths of these arrows by respective ratios a, b, and adding corresponds to putting them together head-to-tail. See Feyman Vol 1 Fig 11-4. Abstractly, the ability to combine vectors in this way is their most fundamental property; if one defined vectors as "objects that can be linearly combined amongst themselves" and unpacked this carefully, one would end up reinventing Definition 8.

1.2 Dot Product

A general vector space only needs enough structure to perform linear combinations, but $\mathbb{R}^n$ has an additional piece of structure in the dot product. Operationally, we compute the dot product of two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ by taking the sum of products of the respective components.

$$\vec{x} \cdot \vec{y} := \sum_{n=1}^{n} x_i y_i.$$

If we endow $\mathbb{R}^n$ with our usual Euclidean notion of distance and angle, then it turns out that this has a nice geometric description.

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \, \|\vec{x}\| \cos(\theta)$$

where $\|\cdot\|$ denotes the length of a vector and where θ is the angle between the two vectors. In particular, for $\vec{x} = \vec{y}$ this reduces to

$$\vec{x} \cdot \vec{x} = \|\vec{x}\|^2$$

which is exactly the n-dimensional version of the Pythagorean theorem. Interestingly, we could tell this story backwards: we could start off with the bare algebraic structure of linear combination, introduce the dot product from thin air, then define length in terms of the dot product

$$\|\vec{x}\| := \sqrt{\vec{x} \cdot \vec{x}}$$

and then take $\vec{x} \cdot \vec{y} = \|\vec{x}\| \, \|\vec{x}\| \cos(\theta)$ as an implicit definition of angles, rather than a derived result. From this point of view, the Pythagorean theorem is demoted to something true by definition!

As a hint of material to come, we notice that from this point of view we could try generalizing the dot product to a so-called inner product which combines $\vec{x}$ with $\vec{y}$ in a more general way to produce a number. For example, we might consider

$$\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot (M\vec{y})$$

for M an $n \times n$ real matrix and see what new notions of distance and angle this yields. If we're not careful with our choice of M, this could give unexpected results. For example, if $M \neq M^T$ we could end up claiming that angle between $\vec{x}$ and $\vec{y}$ is different from the angle between $\vec{y}$ and $\vec{x}$. If M has negative eigenvalues, we will find that some vectors have negative squared lengths

$$\langle \vec{x}, \vec{x} \rangle = \vec{x} \cdot (M\vec{x}) < 0.$$

As strange as this sounds, we will see when we study special relativity that spacetime itself is endowed with such a metric!

1.3   Cross Product

For the special case $n = 3$ we have another way to combine vectors, namely the cross product. There are various equivalent ways to define this, mostly directly by asserting

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} := \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}.$$

For a different approach that more closely resembles the kinds of computations we'll do later in this course, we could also introduce a new object ϵ with three-indexed components ϵ_{ijk} called the Levi-Civita tensor. Just like vectors, tensors have a range of inequivalent definitions at varying levels of abstractions. For the moment, we notice that typographically a vector like $\vec{x} \in \mathbb{R}^n$ has components x_i for i ranging from 1 to n, while a matrix M has 2-indexed components M_{ij}. We can even think of a real number as a single-component object with no indices at all. In this language, a tensor is just a generalization that includes numbers, vectors, and matrices as special cases, but also allows for even more indices. Soon enough we'll give a more precise definition.

Semantics aside, the Levi-Civita tensor ϵ has components ϵ_{ijk} defined recursively by setting

$$\epsilon_{123} = 1$$

and then declaring that ϵ_{ijk} is antisymmetric in its indices, picking up a minus sign whenever we swap any two of them.

$$\epsilon_{ijk} = -\epsilon_{jik} = -\epsilon_{kji} = -\epsilon_{ikj}.$$

Exercise Show that that this uniquely fixes all values of ϵ_{ijk}, and in particular that $\epsilon_{ijk} = 0$ unless i, j, k are all distinct numbers.

Now if we write $\hat{e}_i$ for the ith standard basis vector in $\mathbb{R}^3$ so that, for example,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \hat{e}_1 + x_2 \hat{e}_2 + x_3 \hat{e}_3 = \sum_{i=1}^{3} x_i \hat{e}_i$$

we can equivalently define the cross product by the rule

$$\vec{x} \times \vec{y} = \sum_{i,j=1}^{3} \epsilon_{ijk} x_i y_j \hat{e}_k.$$

Exercise Show that this prescription reproduces the definition of the cross product given earlier.

As it turns out, the cross product too has a neat geometric description: $\vec{x} \times \vec{y}$ is the unique vector such that

- $\vec{x} \times \vec{y}$ is orthogonal to both $\vec{x}$ and $\vec{y}$;
- The length of $\vec{x} \times \vec{y}$ is exactly

$$\|\vec{x} \times \vec{y}\| = \|\vec{x}\| \, \|\vec{y}\| \sin(\theta)$$

  where θ is the angle between $\vec{x}$ and $\vec{y}$; and
- $\vec{x} \times \vec{y}$ points in the direction given by the right hand rule.

This first criterion also tells us why the cross product construction only works nicely in three dimensions - only in $\mathbb{R}^3$ do two non-parallel lines have a unique mutually perpendicular line.

Exercise Show that the cross product acts on basis vectors as

$$\hat{e}_1 \times \hat{e}_2 = \hat{e}_3,$$
$$\hat{e}_2 \times \hat{e}_3 = \hat{e}_1,$$
$$\hat{e}_3 \times \hat{e}_1 = \hat{e}_2.$$

2 Differential Vector Calculus

Now we want to study differential calculus of vectors. What does this mean abstractly? In single-variable calculus we considered functions which took in a single real number and returned a single real number, or abstractly

$$f : \mathbb{R} \to \mathbb{R}.$$

Then differentiating told us things about rates of change, tangent approximations, extrema, and so on. If we want to study vectors using differential calculus, then abstractly we ought to be studying functions that have vectors for their inputs and outputs. These are functions shaped like

$$\vec{F} : \mathbb{R}^n \to \mathbb{R}^m.$$

If we were pure mathematicians, we would proceed by developing definitions and proving theorems that work for totally general values of n and m. But we aren't, so we will concentrate on three cases of classical interest which play inportant roles in the geometry and physics of the Euclidean 3D space which we (to an excellent approximation) inhabit. In the above notation, these are:

1. The trajectory of a particle moving through space as time passes, whose location at time t is given by the vector $\vec{x}(t)$. This takes in a single real number t and return a triple $(x(t), y(t), z(t))$, so it is modeled by a function $\vec{x} : \mathbb{R} \to \mathbb{R}^3$.

2. Scalar fields, which describe things like the electrostatic potential $\phi(\vec{x})$ or the way that temperature $T(\vec{x})$ varies over space at equilibrium. These are modeled by functions $\phi : \mathbb{R}^3 \to \mathbb{R}$.

3. Vector fields, such as the electric field $\vec{E}(\vec{x})$, the magnetic field $\vec{B}(\vec{x})$, or the vector potential $\vec{A}(\vec{x})$ at a fixed time. These take in a position vector $\vec{x}$ and return a vector whose tail we imagine as being based at the location $\vec{x}$, so they are modeled by functions $\vec{A} : \mathbb{R}^3 \to \mathbb{R}^3$

2.1 Trajectories

When we study classical mechanics we learn that if we have a trajectory

$$\vec{x}(t) = (x(t), y(t), z(t))$$

we can differentiate componentwise to find the velocity in Leibniz notation

$$\frac{d\vec{x}}{dt} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) =: (v_x, v_y, v_z)$$

or in Newton flyspeck notation

$$\dot{\vec{x}}(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t)) \, .$$

Differentiating again gives the acceleration

$$\frac{d^2\vec{x}}{dt^2} = \left(\frac{d^2x}{dt^2}, \frac{d^2y}{dt^2}, \frac{d^2z}{dt^2} \right) =: (a_x, a_y, a_z)$$

which gives, among other things, the vector form of Newton's second law if $\vec{x}$ describes the location of a classical point particle subject to a net force $\vec{F}$.

$$\vec{F} = m\frac{d^2\vec{x}}{dt^2}$$

We can reason about how compute the length of the path traversed by our particle in two different but ultimately equivalent and complementary ways. On one hand, in a time dt the particle moves a distance

$$d\ell = (\text{speed})dt$$

where the speed is the norm of the velocity so that

$$d\ell = \|\vec{v}\|\, dt = \sqrt{v_x^2 + v_y^2 + v_z^2}dt.$$

By another route, we could argue that by the 3D version of the Pythagorean theorem, the squared infinitesimal distance must be the sum of the squared infinitesimal changes along each axis

$$d\ell^2 = dx^2 + dy^2 + dz^2.$$

But by the chain we can write in nice suggestive Leibniz notation

$$dx = \frac{dx}{dt}dt$$

and likewise for dy, dz, so that

$$\begin{aligned} d\ell^2 &= dx^2 + dy^2 + dz^2 \\ &= \left(\frac{dx}{dt}dt\right)^2 + \left(\frac{dy}{dt}dt\right)^2 + \left(\frac{dz}{dt}dt\right)^2 \\ &= \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right]dt^2 \\ &= \left(v_x^2 + v_y^2 + v_z^2\right)dt^2 \end{aligned}$$

which is just the square of the result we found before. So the two lines of reasoning agree. In either case, now that we have an expression for $d\ell$ we simply add up all the infinitesimal distances by integration to get the total length of the path. In other words, we have

$$\text{length} = \int_{\text{path}} d\ell = \int_{t_i}^{t_f} \|\vec{v}(t)\|\, dt.$$

Crucially, this is the length of the path traversed by the particle, not the net distance traveled. If a particle moves exactly once around a circle of radius R, it will trace out a path of length $2\pi R$, but having started and ended at the same place it will have moved a net distance of 0.

Exercise Verify this last claim by explicit integration. That is, consider the trajectory

$$\vec{x}(t) = (R\cos(\omega t + a), R\sin(\omega t + a)) \in \mathbb{R}^2$$

for a, ω, R appropriate real constants. Determine a range of t values corresponding to exactly one period, then integrate to find the length of the corresponding path.

5

2.2 Scalar Fields

2.2.1 Gradients and Differentials

We saw that differential calculus tells us about velocity and acceleration vectors (and in principle higher derivatives, though these are less often useful to consider), and that combining this with a bit of integral calculus we can compute things like the length of a path. What about scalar fields $\phi(\vec{x})$? The basic idea of differential calculus is that if we change a function's input a tiny bit its output will change accordingly in a tiny predictable way, so here the basic object we consider is the differential

$$d\phi(\vec{x}) := \phi(\vec{x} + d\vec{x}) - \phi(\vec{x})$$

where $d\vec{x} = (dx, dy, dz)$ is an infinitesimal displacement. In order to describe the differential $d\phi$, we recall the notion of a partial derivative. The partial derivative of $\phi(\vec{x})$ with respect to the first coordinate x, for example, is computed by differentiating $\phi(\vec{x})$ as though it were a function of x alone, treating y and z as constants.

$$\frac{\partial \phi(\vec{x})}{\partial x} := \lim_{h \to 0} \frac{\phi(x+h, y, z) - \phi(x, y, z)}{h}$$

and likewise for the partials with respect to the other components. Then just as the chain rule gives us the nice expression for the differential of a single-variable function

$$df(x) = \frac{df(x)}{dx} dx \text{ for } f : \mathbb{R} \to \mathbb{R}$$

for a scalar field we have

$$d\phi(\vec{x}) = \frac{\partial \phi(\vec{x})}{\partial x} dx + \frac{\partial \phi(\vec{x})}{\partial y} dy + \frac{\partial \phi(\vec{x})}{\partial z} dz \text{ for } \phi : \mathbb{R}^3 \to \mathbb{R}$$

If single-variable functions have tangent line approximations, we can think of this geometrically as a tangent plane expansion. In fact some treatments follow this logic backwards, taking a careful version of this last result as the definition of partial derivatives and reverse-engineering the limit characterization. Now dropping the explicit argument $\vec{x}$ to avoid cluttered notation, we can write this more abstractly as

$$d\phi = \sum_{i=1}^{3} \frac{\partial \phi}{\partial x_i} dx_i \text{ for } \phi : \mathbb{R}^3 \to \mathbb{R}.$$

It is easy to extend this dicussion from $\mathbb{R}^3$ to $\mathbb{R}^n$ - just replace each 3 with an n. We also notice that the last expression above has the form of a dot product. We can rewrite

$$d\phi = \nabla \phi \cdot d\vec{x}$$

where the vector field

$$\nabla \phi := \left(\frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}, \frac{\partial \phi}{\partial z} \right)$$

is called the gradient of ϕ. Some texts write this as simply grad(ϕ). Notice that computing the differential $d\phi$ is totally equivalently to computing the gradient $\nabla \phi$. We have seen gradients before - in electrostatics the electric field is the negative gradient of the electrostatic potential

$$\vec{E} = -\nabla \phi.$$

This in turn owes its namesake to the fact that a particle interacting with a time-independent background potential V experiences a force

$$\vec{F} = -\nabla V.$$

In single-variable calculus we recall that the first derivative tells us how to find extremal points x_* of a function $f(x)$ - just set

$$f'(x_*) = 0 \text{ or equivalently } df(x_*) = 0$$

and solve for x_*. We also recall that this procedure does not tell us whether the resulting values correspond to minima, maxima, or saddle points; for this we have to compute second derivatives. The situation for scalar fields is exactly analogous. The extrema $\vec{x}_*$ of ϕ are given by the condition

$$\nabla\phi(\vec{x}_*) = 0 \text{ or equivalently } d\phi(\vec{x}_*) = 0.$$

In order to distinguish minima, maxima, and the range of intermediate flavors of saddle points, we must consider an object built from second derivatives called the Hessian of ϕ, a symmetric matrix whose entries are all of the mixed second partials of ϕ.

$$H := \begin{bmatrix} \frac{\partial^2 \phi}{\partial x^2} & \frac{\partial^2 \phi}{\partial x \partial y} & \frac{\partial^2 \phi}{\partial x \partial z} \\ \frac{\partial^2 \phi}{\partial y \partial x} & \frac{\partial^2 \phi}{\partial y^2} & \frac{\partial^2 \phi}{\partial y \partial z} \\ \frac{\partial^2 \phi}{\partial z \partial x} & \frac{\partial^2 \phi}{\partial z \partial y} & \frac{\partial^2 \phi}{\partial z^2} \end{bmatrix}$$

We won't get into this here, except to reiterate that this is a *symmetric* matrix. For any nice scalar field[1], the order in which we take repeated partial derivatives does not matter, and in this class we will exclusively consider nice functions.

2.2.2   Chain rule

We recall from single-variable calculus that if f depends on x, while x depends on u, then f depends differentially on u as dictated by the chain rule

$$df = \frac{df}{dx}dx \text{ and } dx = \frac{dx}{du}du \text{ hence } df = \frac{df}{dx}\frac{dx}{du}du$$

The blessing and curse of Leibniz notation is that it makes this hardly look like a theorem at all - just cancel the differentials! Of course this isn't quite true as derivatives aren't really fractions, but it sure makes the theorem easy to remember.

Now what if there are many variables at each layer? Moving to $\mathbb{R}^2$ to avoid cumbersome notation, suppose ϕ depends on $\vec{x} = (x, y)$ while $\vec{x}$ depends on $\vec{u} = (u, v)$. Then from our earlier discussion we have

$$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy$$

while by the same logic

$$dx = \frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv,$$
$$dy = \frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv.$$

Plugging our expressions for dx, dy into our expression for $d\phi$ and organizing our terms gives

$$d\phi = \left(\frac{\partial\phi}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial\phi}{\partial y}\frac{\partial y}{\partial u}\right)du + \left(\frac{\partial\phi}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial\phi}{\partial y}\frac{\partial y}{\partial v}\right)dv.$$

If $\vec{u}$ or $\vec{x}$ had more components the notation would quickly get cumbersome, but if we write

$$\vec{x} = (x_1, \ldots, x_n) \text{ and } \vec{u} = (u_1, \ldots, u_m)$$

then our expression generalizes to

$$d\phi = \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{\partial\phi}{\partial x_i}\frac{\partial x_i}{\partial u_j}du_j.$$

---

[1]For students who know what the following words mean, the exact sufficient condition is that ϕ be continuously differentiable.

Visually, this looks just like the single-variable chain rule but decorated with more indices. Rearranging our terms slightly, we could write this as

$$d\phi = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i} \frac{\partial x_i}{\partial u_j} \right) du_j.$$

Exercise Convince yourself this last manipulation was legal.

If we mildly abuse notation and think of ϕ directly as a function of $\vec{u}$, then we must in the end have

$$d\phi = \sum_{j=1}^{m} \frac{\partial \phi}{\partial u_j} du_j$$

Comparing this last equation to the one before it, we find another way to rephrase the chain rule.

$$\frac{\partial \phi}{\partial u_j} = \sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i} \frac{\partial x_i}{\partial u_j}.$$

As we said this is a mild abuse of notation, since ϕ is only directly a function of $\vec{x}$. Making all arguments explicit gives us the slightly more precise phrasing

$$\frac{\partial \phi(\vec{x}(\vec{u}))}{\partial u_j} = \sum_{i=1}^{n} \left. \frac{\partial \phi(\vec{x})}{\partial x_i} \right|_{\vec{x}=\vec{x}(\vec{u})} \frac{\partial x_i(\vec{u})}{\partial u_j}.$$

This is the analogue of the single-variable formulation of chain rule

$$\frac{d}{du} f(x(u)) = f'(x(u)) x'(u).$$

While this may seem like pedantry, it will be important to be careful about such things as we move forward and abstract out to calculus on manifold.

2.3   Vector Fields

2.3.1   First Derivatives

Now we move to studying the differential calculus of vector fields. We saw that the basic natural operation on a scalar field ϕ is to take its gradient $\nabla \phi$. Abstracting only slightly, we can think of ∇ as a standalone object, a kind of vector whose components are differential operators rather than numbers.

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

But what are the natural ways to differentiate vector fields? We have seen that have several natural ways to turn existing vectors and scalars into new ones: Multiplying a vector and a scalar gives a new vector, taking the dot product of two vectors gives a scalar, and in $\mathbb{R}^3$ taking the cross product of two vectors gives a new vector. These operators on vector and scalars extend naturally to operations on vector fields and scalar fields. For example, we can take the dot product between vector fields $\vec{A}, \vec{B}$ by simply taking their dot products at each point.

$$\left(\vec{A} \cdot \vec{B} \right) (\vec{x}) := \vec{A}(\vec{x}) \cdot \vec{B}(\vec{x})$$

Using the vector operator ∇, this zoo of legal moves also tells us how we can differentiate vector fields.

We have already tried combining the vector ∇ with a scalar to get a new field - this is exactly how we got the gradient. If we take the dot product between ∇ and a vector field $\vec{A}$ we get the divergence

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

and similarly in higher dimensions. Some texts write this as div($\vec{A}$). This operation appears frequently in physics. The first Maxwell equation

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0}.$$

uses it to relate the electric field to the local charge density.

Finally, the cross product between ∇ and a vector field $\vec{A}$ gives the curl, sometimes written as curl($\vec{A}$) or occasionally[2] rot($\vec{A}$)

$$\nabla \times \vec{A} = \begin{bmatrix} \partial_y A_z - \partial_z A_y \\ \partial_z A_x - \partial_x A_z \\ \partial_x A_y - \partial_y A_x \end{bmatrix}$$

where we employ the useful shorthand

$$\partial_x := \frac{\partial}{\partial x}, \partial_z := \frac{\partial}{\partial z}, \partial_z := \frac{\partial}{\partial z}/$$

This operation appears in two of the Maxwell equations

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t},$$

$$\nabla \times \vec{B} = \mu_0 \vec{J} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t}$$
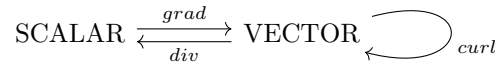
and relates the magnetic field $\vec{B}$ to the vector potential $\vec{A}$ as

$$\vec{B} = \nabla \times \vec{A}.$$

All of the first-order operations div, grad, curl have intuitive physical and geometric meaning (see especially Feynman Vol 2 chapter 2) and could be invented by carefully visualizing fluid flow and electrostatic potentials and so on, but it is interesting to see that we can just as well invent all of it just by enumerating everything our notation permits us to do and trying it!

2.3.2 Second Derivatives

Following the same program, we ought to consider what kinds of second derivatives are legal to make. At a high level, we have three first derivatives div, grad, curl at our disposal. The way these act on vector and scalar fields can be visualized by the following flow chart.

$$\text{SCALAR} \underset{div}{\overset{grad}{\rightleftarrows}} \text{VECTOR} \circlearrowleft {}_{curl}$$

This flow chart tells us all the valid ways to start with either a scalar or a vector field and apply ∇ twice, so we can invent second-order differential calculus for vector by trying every allowed combination. We start with an especially nice and useful legal move, taking the divergence of the gradient of the scalar field. We call this the Laplacian.

$$\nabla \cdot \nabla \phi = \nabla^2 \phi = \sum_i \frac{\partial^2 \phi}{\partial x_i^2}.$$

Combining our earlier equations relating the electrostatic field to the charge density and to the electrostatic potential, we find for example

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} \text{ and } \vec{E} = -\nabla\phi \text{ implies } \nabla^2 \phi = -\frac{\rho}{\epsilon_0}.$$

---

[2]I mention this only because Feynman uses this notion, I haven't seen it anywhere else.

The PDE

$$\nabla^2 \phi(\vec{x}) = f(\vec{x})$$

for f a known function is called Poisson's equation. In the special case $f = 0$ is it the Laplace equation. Both are ubiquitous in theoretical physics. Solutions to these equations are well-studied and have many deep properties, which are responsible for a large number of nice results ranging from fluid mechanics to complex analysis. These are beyond the scope of the present notes, unfortunately.

Next we note a couple of combinations which turn out not to be too interesting, or at least not to correspond to anything geometric and thus tend not to show up directly in physical laws. The first is taking the gradient of the divergence of a vector field.

$$\nabla \left(\nabla \cdot \vec{A} \right) = \text{not too interesting.}$$

This is fine and legal and may come up in the middle of a calculation, but we won't have too much to say about it. Later in this course when we know some differential geometry, we'll have better tools to understand why this doesn't correspond to anything very natural or geometric.

The same is true of taking the curl twice.

$$\nabla \times \left(\nabla \times \vec{A} \right) = \text{not too interesting either.}$$

This doesn't especially lend itself to a neat geometric interpretation, but it is fine to do, and sometimes you'll have to. If you do, it's good to know the following identity

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

where the rightmost term is shorthand for taking the Laplacian of each component A_i separately.

Finally, what happens if we take the curl of a gradient, or the divergence of a curl? As it turns out, both automatically vanish

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

for any $\phi, \vec{A}$, requiring only that these fields are nice enough that the order of partial differentiation doesn't matter (see discussion above).

Exercise Prove that these vanish. This should just be a few lines of algebra.

Now a final aside on vector differential calculus. Rephrasing this last equation slightly: if a vector field happens to be a gradient, then its curl is zero.

$$\vec{A} = \nabla \phi \text{ implies } \nabla \times \vec{A} = 0.$$

The curious reader might wonder if the converse is true. If a vector field has vanishing curl, must it be the gradient of some scalar?

$$\text{Does } \nabla \times \vec{A} = 0 \text{ imply } \vec{A} = \nabla \phi \text{ ?}$$

The answer is yes, and in fact it's not too difficult to cook up an expression for ϕ by hand, but this fact is sensitive to the shape of our space! For example, if we take $\mathbb{R}^3$ and remove the z axis, it is no longer true. There are vector fields on this slightly smaller space with have vanishing curl but which are *not* the gradient of any scalar field. There is an analogous statement for the vanishing of the divergence of the curl.

Physically, this turns out to be a very important fact. It underlies, for example, the Aharonov-Bohm effect in quantum mechanics, and related facts are responsible for a large number of deep phenomena in quantum field theory and condensed matter physics. Mathematically, investigating these facts and their generalizations will lead you down the road of something called de Rham cohomology. Unfortunately all of this is beyond the scope of these notes or even of this course, but it's hard to resist at least mentioning it in passing.

3   Integral Vector Calculus

3.1   Line Integrals

Finally we tread into the land of integral calculus. Our initial intuition will come from the fundamental theorem of single-variable calculus (FTC), which tells us that

$$\int_a^b f'(x)dx = f(b) - f(a).$$

The most immediate generalization to $\mathbb{R}^3$ or $\mathbb{R}^n$ is the line integral. We take a path parametrized as $\vec{x}(t)$ for $t_i \le t \le t_f$ and a vector field $\vec{A}(\vec{x})$, and use this parametrization to define the integral of $\vec{A}$ along the path γ.

$$\int_\gamma \vec{A} \cdot d\vec{\ell} := \int_{t_i}^{t_f} \vec{A}(\vec{x}(t)) \cdot \frac{d\vec{x}(t)}{dt} dt$$

Exercise What if took the same line through $\mathbb{R}^n$ but parametrized it differently? Maybe we use $\vec{y}(t)$ which starts and ends at the same places as $\vec{x}(t)$ and passes through all the same points, but goes more slowly at some parts and more quickly at others. Explain why this should **not** actually change the value of the integral. In other words, the integral we just defined doesn't actually depend on our choice of $\vec{x}(t)$, just on the path it traces out.

If in single-variable calculus we consider the case where our integrand is a total derivative $f'(x)$, here we ought to consider the case of a vector field which happens to be the gradient of some scalar field

$$\vec{A} = \nabla \phi.$$

In that case, for simplicity we write

$$\vec{a} := \vec{x}(t_i) \text{ and } \vec{b} := \vec{x}(t_f).$$

Then the single-variable fundamental theorem of calculus plus the multivariable chain rule tell us

$$\begin{aligned}
\int_\gamma \nabla \phi \cdot d\vec{\ell} &= \int_{t_i}^{t_f} \nabla \phi(\vec{x}(t)) \cdot \frac{d\vec{x}(t)}{dt} dt \\
&= \int_{t_i}^{t_f} \frac{d\phi(\vec{x}(t))}{dt} dt \text{ (chain rule)} \\
&= \phi(\vec{x}(t_f)) - \phi(\vec{x}(t_i)) \text{ (FTC)} \\
&= \phi(\vec{b}) - \phi(\vec{a}).
\end{aligned}$$

This is sometimes called the fundamental theorem of line integrals

3.2   Area and Volume Integrals

The fundamental theorem of line integrals has close generalizations to higher-dimensional integrals. Before we get ahead of ourselves, let's recall how these are defined.

First, consider a two-dimensional surface S, and suppose we make the arbitrary choice[3] of which side of the surface its normal vector $\hat{n}$ ought to point. Then the integral of $\vec{B}$ over the surface S (with the choice of orientation we just made) is just the integral of the normal component of B times an infinitesimal area element dA.

$$\text{Integral of } \vec{B} \text{ over } S = \int_S \vec{B} \cdot \hat{n} dA$$

---

[3]This isn't actually possible for every surface! Look up a diagram of a Klein bottle and convince yourself that it can't be done for that particular shape.

Since this measures how much $\vec{B}$ flows out through S, it is sometimes called the flux of $\vec{B}$.

Now consider a three-dimensional solid region M. We can naturally define the integral of a scalar field over this region, which no need to worry about orientation. We just integrate as though we were computing the volume of M, but with our scalar field as a weighting factor.

$$\text{Integral of } \phi \text{ over } M = \int_M \phi dV$$

For example, if M represented a region of fluid and $\phi(\vec{x})$ represented the density of the fluid at the point $\vec{x}$, this integral would compute the total mass of fluid enclosed in the region.

As we have already mentioned, area and volume integrals obey theorems analogous to the fundamental theorem of line integrals. For example, suppose we consider a two-dimensional surface[4] S with one-dimensional boundary which we write as ∂S. We pick an orientation for the boundary based on the choice of $\hat{n}$ via the right hand rule. If the vector field we integrate over S happens to be a curl, then we get the nice result

$$\int_S \left(\nabla \times \vec{B} \right) \cdot \hat{n} dA = \int_{\partial S} \vec{B} \cdot d\vec{\ell}$$

This is the fundamental theorem of area integrals, sometimes called Stokes' theorem. Some texts call the integral on the RHS the circulation of $\vec{B}$ around ∂S.

Now suppose we take a solid 3D region M with boundary ∂M. If the scalar field we integrate over M happens to be a divergence, then we find

$$\int_M \left(\nabla \cdot \vec{B} \right) \cdot dV = \int_{\partial M} \vec{B} \cdot \hat{n} dA.$$

This is the fundamental theorem of volume integrals, sometimes called the divergence theorem.

As a nice demonstration, we can use these theorems to find equivalent integral formulations of the Maxwell equations. For example, suppose we take Gauss's law for the electric field

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0}.$$

If we integrate both sides over a solid region M with boundary ∂M and apply the divergence theorem, we find

$$\text{Flux of } \vec{E} \text{ through } \partial M = \frac{\text{charge enclosed in } M}{\epsilon_0}.$$

Similarly, if we take Faraday's law

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

and compute the flux of both sides through a surface S with boundary ∂S, we get the equivalent integral form

$$\text{circulation of } \vec{E} \text{ around } \partial S = \frac{d}{dt} \left(\text{flux of } \vec{B} \text{ through } S \right)$$

As it turns out these theorems continue to generalize to something called the generalized Stokes' theorem, written

$$\int_M d\omega = \int_{\partial M} \omega.$$

Here ω is something called a differential form, and $d\omega$ is its so-called exterior derivative, which in the right notation is a kind of general differential object that takes div, grad, curl as special cases. M is likewise some region with boundary ∂M. But we'll need to learn more mathematics before we pursue this abstraction any further.

---

[4]See Feynman Vol 2 Fig 3-7

4   References

This treatment was fairly dry and abstract, and unfortunately had no diagrams or pictures. For a much more intuitive, physics-flavored discussion of mostly the same material see *The Feynman Lectures on Physics* Vol 1 Chapter 11 as well as Vol 2 Chapters 1-3. For similar reasons, see the book *Div, Grad, Curl, and All That* by Schey. To see how div, grad, curl are unified by the notions of exterior derivatives and differential forms, see Chapter 6 of *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach* by Hubbard and Hubbard.