Vector Calculus

University of Cambridge Part IA Mathematical Tripos

David Tong

Department of Applied Mathematics and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge, CB3 OBA, UK

http://www.damtp.cam.ac.uk/user/tong/vc.html
d.tong@damtp.cam.ac.uk

Recommended Books and Resources

There are many good books on vector calculus that will get you up to speed on the basic ideas, illustrated with an abundance of examples.

- H.M Schey, "*Div, Grad, Curl, and all That*"
- Jerrold Marsden and Anthony Tromba, "*Vector Calculus*"

Schey develops vector calculus hand in hand with electromagnetism, using Maxwell's equations as a vehicle to build intuition for differential operators and integrals. Marsden and Tromba is a meatier book but the extra weight is because it goes slower, not further. Neither of these books cover much (if any) material that goes beyond what we do in lectures. In large part this is because the point of vector calculus is to give us tools that we can apply elsewhere and the next steps involve turning to other courses.

- Baxandall and Liebeck, "*Vector Calculus*"

This book does things differently from us, taking a more rigorous and careful path through the subject. For the most part, this involves being more careful from the off about what spaces different objects live in. All of this will be treated in later courses, but if you're someone who likes all their i's dotted, ϵ's small, and $\hbar$'s uncrossed, then this is an excellent place to look.

Contents

Acknowledgements

These lecture notes are far from novel. Large chunks of them have been copied wholesale from the excellent lecture notes of Ben Allanach and Jonathan Evans who previously taught this course. I've also benefitted from the detailed notes of Stephen Cowley. My thanks to Jonathan Evans, Julia Gog and Maria Gutierrez for helpful discussions and suggestions on the material in the course.

I am supported by a Royal Society Wolfson Merit award and by a Simons Investigator award.

0 Introduction

The development of calculus was a watershed moment in the history of mathematics. In its simplest form, we start with a function

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

Provided that the function is continuous and smooth, we can do some interesting things. We can differentiate. And integrate. It's hard to overstate the importance of these operations. It's no coincidence that the discovery of calculus went hand in hand with the beginnings of modern science. It is, among other things, how we describe change.

The purpose of this course is to generalise the concepts of differentiation and integration to functions, or maps, of the form

$$f : \mathbb{R}^m \to \mathbb{R}^n \tag{0.1}$$

with m and n positive integers. Our goal is simply to understand the different ways in which we can differentiate and integrate such functions. Because points in $\mathbb{R}^m$ and $\mathbb{R}^n$ can be viewed as vectors, this subject is called *vector calculus*. It also goes by the name of *multivariable calculus*.

The motivation for extending calculus to maps of the kind (0.1) is manifold. First, given the remarkable depth and utility of ordinary calculus, it seems silly not to explore such an obvious generalisation. As we will see, the effort is not wasted. There are several beautiful mathematical theorems awaiting us, not least a number of important generalisations of the fundamental theorem of calculus to these vector spaces. These ideas provide the foundation for many subsequent developments in mathematics, most notably in geometry. They also underlie every law of physics.

Examples of Maps

To highlight some of the possible applications, here are a few examples of maps (0.1) that we will explore in greater detail as the course progresses. Of particular interest are maps

$$f : \mathbb{R} \to \mathbb{R}^n \tag{0.2}$$

These define *curves* in $\mathbb{R}^n$. A geometer might want to understand how these curves twist and turn in the higher dimensional space or, for $n = 3$, how the curve ties itself in knots. For a physicist, maps of this type are particularly important because they describe the trajectory of a particle. Here the codomain $\mathbb{R}^n$ is identified as physical space, an interpretation that is easiest to sell when $n = 3$ or, for a particle restricted to move on a plane, $n = 2$.

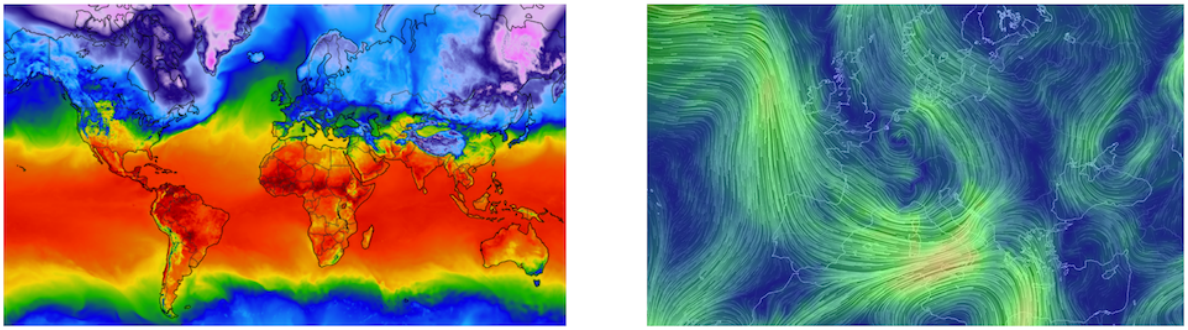


Figure 1. On the left, the temperature on the surface of the Earth is an example of a map from $\mathbb{R}^2 \to \mathbb{R}$, also known as a scalar field. On the right, the wind on the surface of the Earth blows more or less horizontally and so can be viewed as a map from $\mathbb{R}^2 \to \mathbb{R}^2$, also known as a vector field. (To avoid being co-opted by the flat Earth movement, I should mention that, strictly speaking, each of these is a map from $\mathbf{S^2}$ rather than $\mathbb{R}^2$.)

Before we go on, it will be useful to introduce some notation. We'll parameterise $\mathbb{R}$ by the variable t. Meanwhile, we denote points in $\mathbb{R}^n$ as $\mathbf{x}$. A curve (0.2) in $\mathbb{R}^n$ is then written as

$$f : t \to \mathbf{x}(t)$$

Here $\mathbf{x}(t)$ is the image of the map. But, in many situations below, we'll drop the f and just refer to $\mathbf{x}(t)$ as the map. For a physicist, the parameter t is usually viewed as time. In this case, repeated differentiation of the map with respect to t gives us first velocity, and then acceleration.

Going one step further, we could consider maps $f : \mathbb{R}^2 \to \mathbb{R}^n$ as defining a surface in $\mathbb{R}^n$. Again, a geometer might be interested in the curvature of this surface and this, it turns out requires an understanding of how to differentiate the maps. There are then obvious generalisations to higher dimensional surfaces living in higher dimensional spaces.

From the physics perspective, in the map (0.2) that defines a curve the codomain $\mathbb{R}^n$ is viewed as physical space. A conceptually different set of functions arise when we think of the domain $\mathbb{R}^m$ as physical space. For example, we could consider maps of the kind

$$f : \mathbb{R}^3 \to \mathbb{R}$$

where $\mathbb{R}^3$ is viewed as physical space. Physicists refer to this as a *scalar field*. (Mathematicians refer to it as a map from $\mathbb{R}^3$ to $\mathbb{R}$.) A familiar example of such a map is temperature: there exists a temperature at every point in this room and that gives a map $T(\mathbf{x})$. This is shown in Figure 1. A more fundamental, and ultimately more interesting, example of a scalar field is the Higgs field in the Standard Model of particle physics.

As one final example, consider maps of the form

$$f : \mathbb{R}^3 \to \mathbb{R}^3$$

where, again, the domain $\mathbb{R}^3$ is identified with physical space. Physicists call these *vector fields*. (By now, you can guess what mathematicians call them.) In fundamental physics, two important examples are provided by the electric field $\mathbf{E}(\mathbf{x})$ and magnetic field $\mathbf{B}(\mathbf{x})$, first postulated by Michael Faraday: each describes a three-dimensional vector associated to each point in space.

1 Curves

In this section, we consider maps of the form

$$f : \mathbb{R} \to \mathbb{R}^n$$

A map of this kind is called a *parameterised curve*, with $\mathbb{R}$ the parameter and the curve the image of the map in $\mathbb{R}^n$. In what follows, we will denote the curve as C.

Whenever we do explicit calculations, we need to introduce some coordinates. The obvious ones are Cartesian coordinates, in which the vector $\mathbf{x} \in \mathbb{R}^n$ is written as

$$\mathbf{x} = (x^1, \ldots, x^n) = x^i \mathbf{e}_i$$

where, in the second expression, we're using summation convention and explicitly summing over $i = 1, \ldots, n$. Here $\{\mathbf{e}_i\}$ is a choice of orthonormal basis vectors, satisfying $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. For $\mathbb{R}^n = \mathbb{R}^3$, we'll also write these as $\{\mathbf{e}_i\} = \{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$. (The notation $\{\mathbf{e}_i\} = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is also standard, although we won't adopt it in these lectures.)

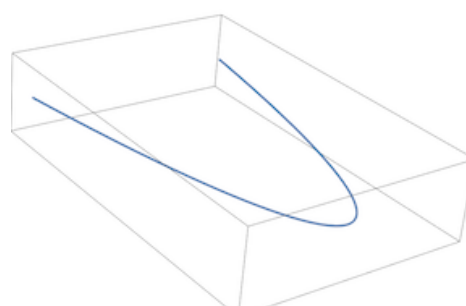
The image of the function can then be written as $\mathbf{x}(t)$. In physics, we might think of this as the trajectory of a particle evolving in time t. Here, we'll mostly just view the curve as an abstract mathematical map, with t nothing more than a parameter labelling positions along the curve. In fact, one of themes of this section is that, for many calculations, the choice of parameter t is irrelevant.

Examples

Here are two simple examples. Consider first the map $\mathbb{R} \to \mathbb{R}^3$ that takes the form

$$\mathbf{x}(t) = (at, bt^2, 0)$$

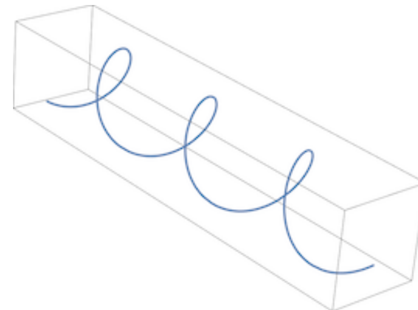
The image of the map is the parabola $a^2 y = bx^2$, lying in the plane $z = 0$, and is shown on the right.



This looks very similar to what you would draw if asked to plot the graph $y = bx^2/a^2$, with the additional requirement of $z = 0$ prompting the artistic flourish that results in a curve suspended in 3d. Obviously, the curve $\mathbf{x}(t)$ and the functions $y = bx^2/a^2$ (with $z = 0$) are related, but they're not quite the same thing. The function $y = bx^2/a^2$ is usually thought of as a map $\mathbb{R} \to \mathbb{R}$ and in plotting a graph you include both the domain and codomain. In contrast, on the right we've plotted only the image of the curve $\mathbf{x}(t)$ in $\mathbb{R}^3$; the picture loses all information about the domain coordinate t.

Here is a second example that illustrates the same point. Consider

$$\mathbf{x}(t) = (\cos t, \sin t, t) \tag{1.1}$$



The resulting curve is a helix, shown to the right. Like any other curve, the choice of parameterisation is not unique. We could, for example, consider the different map

$$\mathbf{x}(t) = (\cos \lambda t, \sin \lambda t, \lambda t)$$

This gives the same helix as (1.1) for any choice of $\lambda \in \mathbb{R}$ as long as $\lambda \neq 0$. In some contexts this matters. If, for example, t is time, and $\mathbf{x}(t)$ is the trajectory of a rollercoaster then the fate of the contents of your stomach depends delicately on the value of λ. However, there will be some properties of the curve that are independent of the choice of parameterisation and, in this example, independent of λ. It is these properties that will be our primary interest in this section.

Before we go on, a pedantic mathematical caveat. It may be that the domain of the curve is not all of $\mathbb{R}$. For example, we could have the map $\mathbb{R} \to \mathbb{R}^2$ given by $\mathbf{x}(t) = (t, \sqrt{1-t^2})$. This makes sense only for the interval $t \in [-1, +1]$ and you should proceed accordingly.

1.1 Differentiating the Curve

The vector function $\mathbf{x}(t)$ is *differentiable* at t if, as $\delta t \to 0$, we can write

$$\mathbf{x}(t + \delta t) - \mathbf{x}(t) = \dot{\mathbf{x}}(t)\,\delta t + \mathcal{O}(\delta t^2) \tag{1.2}$$

You should think of this expression as defining the derivative $\dot{\mathbf{x}}(t)$. If the derivative $\dot{\mathbf{x}}$ exists everywhere then the curve is said to be *smooth*. This means that it is continuous and, as the name suggests, not egregiously jagged.

There are some notational issues to unpick in this expression. First, $\mathcal{O}(\delta t^2)$ includes all terms that scale as δt^2 or smaller as $\delta t \to 0$. This "big-O" notation is commonly used in physics and applied mathematics. In pure maths you will also see the "little o" notation $o(\delta t)$ which means "strictly smaller than δt" as $\delta t \to 0$. Roughly speaking $o(\delta t)$ is the same thing as $\mathcal{O}(\delta t^2)$. (In other courses you may encounter situations where this speaking is too rough to be accurate, but it will suffice for our needs.) We'll stick with big-O notation throughout these lectures.

We've denoted the derivative in (1.2) with a dot, $\dot{\mathbf{x}}(t)$. This was Newton's original notation for the derivative and, 350 years later, comes with some sociological baggage. In physics, a dot is nearly always used to denote differentiation with respect to time, so the velocity of a particle is $\dot{\mathbf{x}}$ and the acceleration is $\ddot{\mathbf{x}}$. Meanwhile a prime, like $f'(x)$, is usually used to denote differentiation with respect to space. This is deeply ingrained in the psyche of physicists, so much so that I get a little shudder if I see something like $x'(t)$, even though it's perfectly obvious that it means dx/dt. Mathematicians, meanwhile, seem to have no such cultural hang-ups on this issue. (They reserve their cultural hang-ups for a 1000 other issues.)

We write the left-hand side of (1.2) as

$$\delta\mathbf{x}(t) = \mathbf{x}(t+\delta t) - \mathbf{x}(t)$$

The derivative is then the vector

$$\frac{d\mathbf{x}}{dt} = \dot{\mathbf{x}}(t) = \lim_{\delta t\to 0}\frac{\delta\mathbf{x}}{\delta t}$$

Here the familiar notation $d\mathbf{x}/dt$ for the derivative is due to Leibniz and works if we're differentiating with respect to time, space, or anything else. We'll also sometimes use the slightly sloppy notation and write

$$d\mathbf{x} = \dot{\mathbf{x}}\,dt$$

which, at least for now, really just means the same thing as (1.2) except we've dropped the $\mathcal{O}(\delta t^2)$ terms.

It's not difficult to differentiate vectors and, at least in Cartesian coordinates with the basis vectors $\mathbf{e}_i$, we can just do it component by component

$$\mathbf{x}(t) = x^i(t)\mathbf{e}_i \quad\Rightarrow\quad \dot{\mathbf{x}}(t) = \dot{x}^i(t)\mathbf{e}_i$$

The same is true if we work in any other choice of basis vectors $\{\mathbf{e}_i\}$ provided that these vectors themselves are independent of t. (In the lectures on Dynamics and Relativity we encounter an example where the basis vectors *do* depend on time and you have to be more careful. This arises in Section 6 on "Non-Inertial Frames".)

More generally, given a function $f(t)$ and two vector functions $\mathbf{g}(t)$ and $\mathbf{h}(t)$, it's simple to check that the following Leibniz identities hold

$$\frac{d}{dt}(f\mathbf{g}) = \frac{df}{dt}\mathbf{g} + f\frac{d\mathbf{g}}{dt}$$
$$\frac{d}{dt}(\mathbf{g}\cdot\mathbf{h}) = \frac{d\mathbf{g}}{dt}\cdot\mathbf{h} + \mathbf{g}\cdot\frac{d\mathbf{h}}{dt}$$

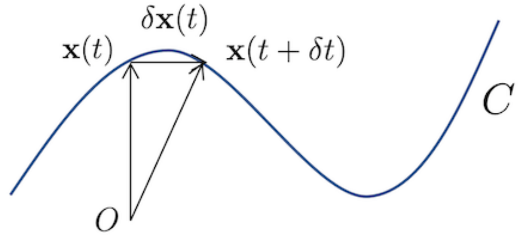


Figure 2. The derivative is the tangent vector to the curve.

Moreover, if $\mathbf{g}(t)$ and $\mathbf{h}(t)$ are vectors in $\mathbb{R}^3$, we also have the cross-product identity

$$\frac{d}{dt}(\mathbf{g} \times \mathbf{h}) = \frac{d\mathbf{g}}{dt} \times \mathbf{h} + \mathbf{g} \times \frac{d\mathbf{h}}{dt}$$

As usual, we have to be careful with the ordering of terms in the cross product because for example, $d\mathbf{g}/dt \times \mathbf{h} = -\mathbf{h} \times d\mathbf{g}/dt$.

1.1.1 Tangent Vectors

There is a nice geometric meaning to the derivative $\dot{\mathbf{x}}(t)$ of a parameterised curve C: it gives the tangent to the curve and is called, quite reasonably, the *tangent vector*. This is shown in Figure 2.

The direction of the tangent vector $\dot{\mathbf{x}}(t)$ is geometrical (at least up to a sign): it depends only on the curve C itself, and not on the choice of parameterisation. In contrast, the magnitude of the tangent vector $|\dot{\mathbf{x}}(t)|$ does depend on the parameterisation. This is obvious mathematically, since we're differentiating with respect to t, and also physically where $\dot{\mathbf{x}}$ is identified with the velocity of a particle.

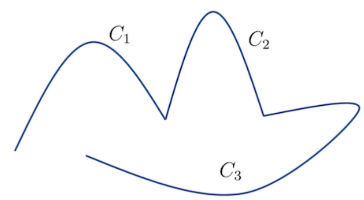
Sometimes, you may find yourself with an unwise choice of parameterisation in which the derivative vector $\dot{\mathbf{x}}$ vanishes at some point. For example, consider the curve in $\mathbb{R}^2$ given by

$$\mathbf{x}(t) = (t^3, t^3)$$

The curve C is just the straight line $x = y$. The tangent vector $\dot{x} = 3t^2(1,1)$ which clearly points along the line $x = y$ but with magnitude $3\sqrt{2}t^2$ and so vanishes at $t = 0$. Clearly this is not a property of C itself, but of our choice of parameterisation. We get the same curve C from the map $\mathbf{x}(t) = (t, t)$ but now the tangent vector is everywhere non-vanishing.

A parameterisation is called *regular* if $\dot{\mathbf{x}}(t) \neq 0$ for any t. In what follows, we will assume that we are dealing with regular parameterisations except, perhaps, at isolated points. This means that we can divide the curve into segments, each of which is regular.

As a slightly technical aside, we will sometimes have cause to consider curves that are *piecewise smooth* curves of the form $C = C_1 + C_2 + \ldots$, where the end of one curve lines up with the beginning of the next, as shown on the right. In this case, a tangent vector exists everywhere except at the cusps where two curves meet.



1.1.2 The Arc Length

We can use the tangent vectors to compute the length of the curve. From Figure 2, we see that the distance between two nearby points is

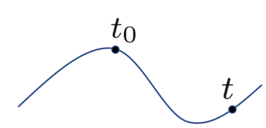
$$\delta s = |\delta \mathbf{x}| + \mathcal{O}(|\delta \mathbf{x}|^2) = |\dot{\mathbf{x}}\, \delta t| + \mathcal{O}(\delta t^2)$$

We then have

$$\frac{ds}{dt} = \pm \left| \frac{d\mathbf{x}}{dt} \right| = \pm |\dot{\mathbf{x}}| \tag{1.3}$$

where we get the plus sign for distances measured in the direction of increasing t, and the minus sign in the direction of decreasing t.

If we pick some starting point t_0 on the curve, then the distance along the curve to any point $t > t_0$ is given by



$$s = \int_{t_0}^{t} dt' \; |\dot{\mathbf{x}}(t')|$$

This distance is called the *arc length*, s. Because $|\dot{\mathbf{x}}| > 0$, this is a positive and strictly increasing function as we move away in the direction $t > t_0$. It is a negative, and strictly decreasing function in the direction $t < t_0$.

Although the tangent vector $\dot{\mathbf{x}}$ depends on the choice of parameterisation, the arc length s does not. We can pick a different parameterisation of the curve $\tau(t)$, which we will take to be an invertible and smooth function. We will also assume that $d\tau/dt > 0$

so that they both measure "increasing time" in the same direction. The chain rule tells us that

$$\frac{d\mathbf{x}}{dt} = \frac{d\mathbf{x}}{d\tau}\frac{d\tau}{dt} \tag{1.4}$$

We can then compute the arc length using the τ parameterisation: it is

$$s = \int_{t_0}^{t} dt'\ |\dot{\mathbf{x}}(t')| = \int_{\tau_0}^{\tau} d\tau'\ \frac{dt'}{d\tau'}\left|\frac{d\mathbf{x}}{d\tau'}\frac{d\tau'}{dt'}\right| = \int_{\tau_0}^{\tau} d\tau'\ \left|\frac{d\mathbf{x}}{d\tau'}\right| \tag{1.5}$$

In the second equality, we find the contribution from the chain rule (1.4) together with a factor from the measure that comes from integrating over $d\tau$ instead of dt. These then cancel in the third equality. The upshot is that we can compute the arc length using any parameterisation that we wish. Or, said differently, the arc length is independent of the choice of parameterisation of the curve.

We can now turn this on its head. All parameterisations of the curve give the same arc length. But this means that the arc length itself is, in many ways, the only natural parameterisation of the curve. We can then think of $\mathbf{x}(s)$ with the corresponding tangent vector $d\mathbf{x}/ds$. From (1.3), we see that this choice of the tangent vector always has unit length: $|d\mathbf{x}/ds| = 1$.

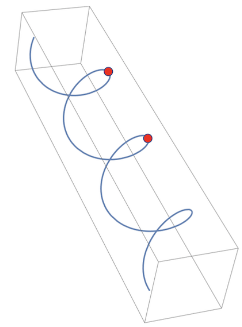
As an aside: these kind of issues raise their head in the physics of special relativity where time means different things for people moving at different speeds. This means that there is no universally agreed "absolute time" and so different people will parameterise the trajectory of a particle $\mathbf{x}(t)$ in different ways. There's no right or wrong way, but it's annoying if someone does it differently to you. (Admittedly, this is only likely to happen if they are travelling at an appreciable fraction of the speed of light relative to you.) Happily there is something that everyone can agree on, which is the special relativistic version of arc length. It's known as *proper time*. You can read more about this in the lectures on Dynamics and Relativity.

An Example

To illustrate these ideas, let's return to our helix example of (1.1). We had $\mathbf{x}(t) = (\cos t, \sin t, t)$ and so $\dot{\mathbf{x}}(t) = (-\sin t, \cos t, 1)$. Our defining equation (1.3) then becomes (taking the positive sign)

$$\frac{ds}{dt} = |\dot{\mathbf{x}}| = \sqrt{2}$$

If we take $t_0 = 0$, then the arc length measured from the point $\mathbf{x} = (1, 0, 0)$ is $s = \sqrt{2}t$. In particular, after time $t = 2\pi$ we've made a full rotation and sit at $\mathbf{x} = (1, 0, 2\pi)$. These two points are shown as red dots in the figure. Obviously the direct route between the two has distance 2π. Our analysis above shows that the distance along the helix is $s = \sqrt{8}\pi$.



1.1.3 Curvature and Torsion

There is a little bit of simple geometry associated to these ideas. Given a curve C, parameterised by its arc length s, we have already seen that the tangent vector

$$\mathbf{t} = \frac{d\mathbf{x}}{ds}$$

has unit length, $|\mathbf{t}| = 1$. (Note: don't confuse the bold faced tangent vector $\mathbf{t}$ with our earlier parameterisation t: they're different objects!) We can also consider the "acceleration" of the curve with respect to the arc length, $d^2\mathbf{x}/ds^2$. The magnitude of this "acceleration" is called the *curvature*

$$\kappa(s) = \left|\frac{d^2\mathbf{x}}{ds^2}\right| \tag{1.6}$$

To build some intuition, we can calculate the curvature of a circle of radius R. If we start with a simple parameterisation $\mathbf{x}(t) = (R\cos t, R\sin t)$ then you can check using (1.3) that the arc length is $s = Rt$. We then pick the new parameterisation $\mathbf{x}(s) = (R\cos(s/R), R\sin(s/R))$. We then find that a circle of radius R has constant curvature

$$\kappa = \frac{1}{R}$$

Note, in particular, that as $R \to \infty$, the circle becomes a straight line which has vanishing curvature.

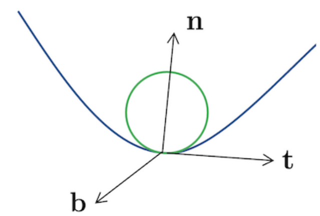
There is also a unit vector associated to this "acceleration", defined as

$$\mathbf{n} = \frac{1}{\kappa}\frac{d^2\mathbf{x}}{ds^2} = \frac{1}{\kappa}\frac{d\mathbf{t}}{ds}$$

This is known as the *principal normal*. Note that the factor of $1/\kappa$ ensures that $|\mathbf{n}| = 1$.

Importantly, if $\kappa \neq 0$ then $\mathbf{n}$ is perpendicular to the tangent vector $\mathbf{t}$. This follows from the fact that $\mathbf{t} \cdot \mathbf{t} = 1$ and so $d/ds(\mathbf{t} \cdot \mathbf{t}) = 2\kappa \mathbf{n} \cdot \mathbf{t} = 0$. This means that $\mathbf{t}$ and $\mathbf{n}$ define a plane, associated to every point in the curve. This is known as the *osculating plane*.

For any point s on the curve, there is an associated osculating plane. Now draw a circle in that plane that touches the curve at the point s, whose curvature matches $\kappa(s)$. This is called the *osculating circle* and is shown in green in the figure. This is the circle that just kisses the curve at s



Next we can ask: how does the osculating plane vary as we move along the curve? This is simplest to discuss if we restrict to curves in $\mathbf{R}^3$. In this case, we have the cross product at our disposal and we can define the unit normal to the osculating plane as

$$\mathbf{b} = \mathbf{t} \times \mathbf{n}$$

This is known as the *binormal*, to distinguish it from the normal $\mathbf{n}$. The three vectors $\mathbf{t}$, $\mathbf{n}$ and $\mathbf{b}$ define an orthonormal basis for $\mathbb{R}^3$ at each point s along the curve (at least as long as $\kappa(s) \neq 0$.) This basis twists and turns along the curve.

Note that $|\mathbf{b}| = 1$ which, using the same argument as for $\mathbf{t}$ above, tells us that $\mathbf{b} \cdot d\mathbf{b}/ds = 0$. In addition, we have $\mathbf{t} \cdot \mathbf{b} = 0$ which, after differentiating, tells us that

$$0 = \frac{d\mathbf{t}}{ds} \cdot \mathbf{b} + \mathbf{t} \cdot \frac{d\mathbf{b}}{ds} = \kappa \mathbf{n} \cdot \mathbf{b} + \mathbf{t} \cdot \frac{d\mathbf{b}}{ds}$$

But, by definition, $\mathbf{n} \cdot \mathbf{b} = 0$. So we learn that $\mathbf{t} \cdot d\mathbf{b}/ds = 0$. In other words, $d\mathbf{b}/ds$ is orthogonal to both $\mathbf{b}$ and to $\mathbf{t}$. Which means that it must be parallel to $\mathbf{n}$. We define the *torsion* $\tau(s)$ as a measure of how the binormal changes

$$\frac{d\mathbf{b}}{ds} = -\tau(s)\mathbf{n} \tag{1.7}$$

From the definition, you can see that the torsion is a measure of $\dddot{\mathbf{x}}$. The minus sign means that if the top of the green circle in the figure tilts towards us, then $\tau > 0$; if it tilts away from us then $\tau < 0$. Heuristically, the curvature captures how much the curve fails to be a straight line, while the torsion captures how much the curve fails to be planar.

The Frenet-Serret Equations

There is a closed set of formulae describing curvature and torsion. These are the *Frenet-Serret* equations,

$$\frac{d\mathbf{t}}{ds} = \kappa\mathbf{n} \tag{1.8}$$
$$\frac{d\mathbf{b}}{ds} = -\tau\mathbf{n} \tag{1.9}$$
$$\frac{d\mathbf{n}}{ds} = \tau\mathbf{b} - \kappa\mathbf{t} \tag{1.10}$$

The first of these (1.8) is simply the definition of the normal $\mathbf{n}$.

That leaves us with (1.10). We'll again start with the definition $\mathbf{b} = \mathbf{t} \times \mathbf{n}$, and this time take the cross product with $\mathbf{t}$. The triple product formula then gives us

$$\mathbf{b} \times \mathbf{t} = (\mathbf{t} \times \mathbf{n}) \times \mathbf{t} = (\mathbf{t} \cdot \mathbf{t})\,\mathbf{n} - (\mathbf{n} \cdot \mathbf{t})\,\mathbf{t} = \mathbf{n}$$

Now taking the derivative with respect to s, using (1.8) and (1.9) and noting that $\mathbf{b} = \mathbf{t} \times \mathbf{n}$ and $\mathbf{t} = \mathbf{n} \times \mathbf{b}$ then gives us (1.10).

It's useful to rewrite the first two equations (1.8) and (1.9) using $\mathbf{n} = \mathbf{b} \times \mathbf{t}$ so that we have

$$\frac{d\mathbf{t}}{ds} = \kappa(\mathbf{b} \times \mathbf{t}) \quad \text{and} \quad \frac{d\mathbf{b}}{ds} = -\tau(\mathbf{b} \times \mathbf{t})$$

This is six first order equations for six unknowns, $\mathbf{b}(s)$ and $\mathbf{t}(s)$. If we are given $\kappa(s)$ and $\tau(s)$, together with initial conditions $\mathbf{b}(0)$ and $\mathbf{t}(0)$, then we can solve for $\mathbf{b}(s)$ and $\mathbf{t}(s)$ and can subsequently solve for the curve $\mathbf{x}(s)$. The way to think about this is that the curvature and torsion $\kappa(s)$ and $\tau(s)$ specify the curve, up to translation and orientation.

1.2 Line Integrals

Given a curve C in $\mathbb{R}^n$ and some function defined over $\mathbb{R}^n$, we may well wish to integrate the function along the curve. There are different stories to tell for scalar and vector fields and we deal with each in turn.

1.2.1 Scalar Fields

A scalar field is a map

$$\phi : \mathbb{R}^n \to \mathbb{R}$$

With coordinates $\mathbf{x}$ on $\mathbb{R}^n$, we'll denote this scalar field as $\phi(\mathbf{x})$.

Given a parameterised curve C in $\mathbb{R}^n$, which we denote as $\mathbf{x}(t)$, it might be tempting to put these together to get the function $\phi(\mathbf{x}(t))$ which is a composite map $\mathbb{R} \to \mathbb{R}$. We could then just integrate over t in the usual way.

However, there's a catch. The result that you get will depend both on the function ϕ, the curve C, *and* the choice of parameterisation of the curve. There's nothing wrong this per se, but it's not what we want here. For many purposes, it turns out to be more useful to have a definition of the integral that depends only on the function ϕ and the curve C, but gives the same answer for any choice of parameterisation of the curve.

One way to achieve this is to work with the arc length s which, as we've seen, is the natural parameterisation along the curve. We can integrate from point $\mathbf{a}$ to point $\mathbf{b}$, with $\mathbf{x}(s_a) = \mathbf{a}$ and $\mathbf{x}(s_b) = \mathbf{b}$ and $s_a < s_b$, by defining the *line integral*

$$\int_C \phi \; ds = \int_{s_a}^{s_b} \phi(\mathbf{x}(s)) \; ds$$

where the right-hand side is now viewed as a usual one-dimensional integral.

This line integral is, by convention, defined so that $\int_C ds$ gives the length of the curve C and, in particular, is always positive. In other words, there's no directional information in this integral: it doesn't matter what way you move along the curve.

Suppose that we're given a parameterised curve C in terms of some other parameter $\mathbf{x}(t)$, with $\mathbf{x}(t_a) = \mathbf{a}$ and $\mathbf{x}(t_b) = \mathbf{b}$. The usual change of variables tells us that

$$\int_C \phi \; ds = \int_{t_a}^{t_b} \phi(\mathbf{x}(t)) \, \frac{ds}{dt} \, dt$$

We can then use (1.3). If $t_b > t_a$ then we have $ds/dt = +|\dot{\mathbf{x}}|$ and

$$\int_C \phi \; ds = \int_{t_a}^{t_b} \phi(\mathbf{x}(t)) \, |\dot{\mathbf{x}}(t)| \; dt \tag{1.11}$$

Meanwhile, if $t_b < t_a$ then we have $ds/dt = -|\dot{\mathbf{x}}|$ and

$$\int_C \phi \; ds = \int_{t_b}^{t_a} \phi(\mathbf{x}(t)) \, |\dot{\mathbf{x}}(t)| \; dt$$

We see that the line integral comes with the length of the tangent vector $|\dot{\mathbf{x}}|$ in the integrand. This is what ensures that the line integral is actually independent of the choice of parameterisation: the argument is the same as the one we used in (1.5) to show that the arc length is invariant under reparameterisations: upon a change of variables, the single derivative d/dt in $\dot{\mathbf{x}}$ cancels the Jacobian from the integral $\int dt$. Furthermore, the minus signs work out so that you're always integrating from a smaller value of t to a larger one, again ensuring that $\int_C ds$ is positive and so can be interpreted as the length of the curve.

1.2.2 Vector Fields

Vector fields are maps of the form

$$\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$$

So that at each point $\mathbf{x} \in \mathbb{R}^n$ we have a vector-valued object $\mathbf{F}(\mathbf{x})$. We would like to understand how to integrate a vector field along a curve C.

There are two ways to do this. We could work component-wise, treating each component like the scalar field example above. After doing the integration, this would leave us with a vector.

However it turns out that, in many circumstances, it's more useful to integrate the vector field so that the integral gives us just a number. We do this integrating the component of the vector field that lies tangent to the curve. Usually, this is what is meant by the *line integral* of a vector field.

In more detail, suppose that our curve C has a parameterisation $\mathbf{x}(t)$ and we wish to integrate from t_a to t_b, with $\mathbf{x}(t_a) = \mathbf{a}$ and $\mathbf{x}(t_b) = \mathbf{b}$. The line integral of a vector field $\mathbf{F}$ along C is defined to be

$$\int_C \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_{t_a}^{t_b} \mathbf{F}(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t)\ dt \tag{1.12}$$

Once again, this doesn't depend on the choice of parameterisation t. This is manifest in the expression on the left where the parameterisation isn't mentioned. The right-hand side is invariant for the same reason as (1.11).

This time, however, there's a slightly different story to tell about minus signs. We should think of each curve C as coming with an *orientation*, which is the direction along the curve. Equivalently, it can be thought of as the direction of the tangent vector $\dot{\mathbf{x}}$. In the example above, the orientation of the curve is from $\mathbf{a}$ to $\mathbf{b}$. This then determines the limits of the integral, from t_a to t_b, since $\mathbf{x}(t_a) = \mathbf{a}$ and $\mathbf{x}(t_b) = \mathbf{b}$. Note that the limits are always this way round, regardless of whether our parameterisation has $t_a < t_b$ or whether $t_b > t_a$: the orientation determines the limits, not the parameterisation.

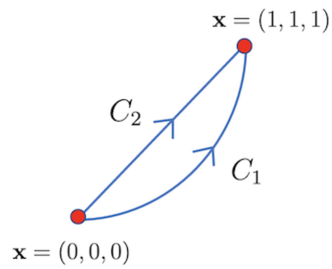
In summary, the line integral for a scalar field $\int_C \phi\, ds$ is independent of the orientation and, if ϕ is positive, the integral will also be positive. In contrast, the integral of the vector field $\int_C \mathbf{F}\cdot\dot{\mathbf{x}}\, dt$ depends on the orientation. Flip the orientation of the curve, and the integral will change sign.

An Example

As a slightly baroque example, consider the vector field in $\mathbb{R}^3$,

$$\mathbf{F}(\mathbf{x}) = (xe^y, z^2, xy)$$

To evaluate the line integral, we also need to specify the curve C along which we perform the integral. We'll consider two options, both of which evolve from $\mathbf{x}(t=0) = (0,0,0)$ to $\mathbf{x}(t=1) = (1,1,1)$. Our first curve is



$$C_1: \qquad \mathbf{x}(t) = (t, t^2, t^3)$$

This is shown in the figure. Evaluated on C_1, we have $\mathbf{F}(\mathbf{x}(t)) = (te^{t^2}, t^6, t^3)$. Meanwhile $\dot{\mathbf{x}} = (1, 2t, 3t^2)$ so we have

$$\begin{aligned}\int_{C_1} \mathbf{F}\cdot d\mathbf{x} &= \int_0^1 dt\ \mathbf{F}\cdot\dot{\mathbf{x}} \\ &= \int_0^1 dt\ \left(te^{t^2} + 2t^7 + 3t^5\right) = \frac{1}{4}(1+2e)\end{aligned}$$

Our second curve is simply the straight line

$$C_2: \qquad \mathbf{x}(t) = (t, t, t)$$

Evaluated on this curve, we have $\mathbf{F}(\mathbf{x}(t)) = (te^t, t^2, t^2)$. Now the tangent vector is $\dot{x} = (1,1,1)$ and the integral is

$$\int_{C_2} \mathbf{F}\cdot d\mathbf{x} = \int_0^1 dt\ \mathbf{F}\cdot\dot{\mathbf{x}} = \int_0^1 dt\ \left(te^t + 2t^2\right) = \frac{5}{3} \tag{1.13}$$

(The first of these integrals is done by an integration by parts.)

The main lesson to take from this is the obvious one: the answers are different. The result of a line integral generally depends on both the thing you're integrating $\mathbf{F}$ and the choice of curve C.

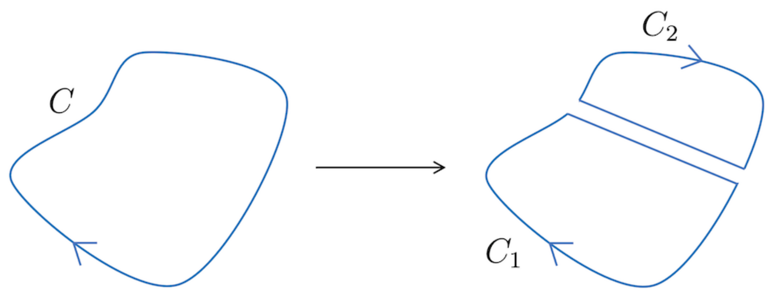


Figure 3. Decomposing a curve by introducing new segments with opposite orientations.

More Curves, More Integrals

We'll see plenty more examples of line integrals, both in this course and in later ones. Here are some comments to set the scene.

First, there will be occasions when we want to perform a line integral around a *closed curve* C, meaning that the starting and end points are the same, $\mathbf{a} = \mathbf{b}$. For such curves, we introduce new notation and write the line integral as

$$\oint_C \mathbf{F} \cdot d\mathbf{x}$$

with the little circle on the integral sign there to remind us that we're integrating around a loop. This quantity is called the *circulation* of $\mathbf{F}$ around C. The name comes from Fluid Mechanics where we might view $\mathbf{F}$ as the velocity field of a fluid, and the circulation quantifies the swirling motion of the fluid.

In other occasions, we may find ourselves in a situation in which the curve C decomposes into a number of piecewise smooth curves C_i, joined up at their end points. We write $C = C_1 + C_2 + \ldots$, and the line integral is

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_{C_1} \mathbf{F} \cdot d\mathbf{x} + \int_{C_2} \mathbf{F} \cdot d\mathbf{x} + \ldots$$

It is also useful to think of the curve $-C$ as the same as the curve C but with the opposite orientation. This means that we have the expression

$$\int_{-C} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = -\int_C \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x}$$

For example, we could return to our previous baroque example and consider the closed curve $C = C_1 - C_2$. This curve starts at $\mathbf{x} = (0,0,0)$, travels along C_1 to $\mathbf{x} = (1,1,1)$

and then returns back along C_2 in the opposite direction to the arrow. From our previous answers, we have

$$\oint_C \mathbf{F} \cdot d\mathbf{x} = \int_{C_1} \mathbf{F} \cdot d\mathbf{x} - \int_{C_2} \mathbf{F} \cdot d\mathbf{x} = \frac{1}{4}(1 + 2e) - \frac{5}{3}$$

There are lots of games that we can play like this. For example, it's sometimes useful to take a smooth closed curve C and decompose it into two piecewise smooth segments,. An example is shown in Figure 3, where we've introduced two new segments, which should be viewed as infinitesimally close to each other. These two new segments have opposite orientation and so cancel out in any integral. In this way, we can think of the original curve as $C = C_1 + C_2$. We'll see other examples of these kinds of manipulations as we progress.

1.3 Conservative Fields

Here's an interesting question. In general the line integral of a vector field depends on the path taken. But is this ever not the case? In other words, are there some vector fields $\mathbf{F}$ for which the line integral depends only on the end points and not on the route you choose to go between them?

Such a vector field $\mathbf{F}$ would obey

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{x} = \int_{C_2} \mathbf{F} \cdot d\mathbf{x}$$

for any C_1 and C_2 that share the same end points $\mathbf{a}$ and $\mathbf{b}$ and the same orientation. Equivalently, we could consider the closed curve $C = C_1 - C_2$ and write this as

$$\oint_C \mathbf{F} \cdot d\mathbf{x} = 0$$

for all closed curves C. To answer this question about vector fields, we first need to introduce a new concept for scalar fields.

1.3.1 The Gradient

Let's return to the scalar field

$$\phi : \mathbb{R}^n \to \mathbb{R}$$

We want to ask: how can we differentiate such a function?

With Cartesian coordinates $\mathbf{x} = (x^1, \ldots, x^n)$ on $\mathbb{R}^n$, the scalar field is a function $\phi(x^1, \ldots, x^n)$. Given such a function of several variables, we can always take *partial derivatives*, which means that we differentiate with respect to one variable while keeping all others fixed. For example,

$$\frac{\partial \phi}{\partial x^1} = \lim_{\epsilon \to 0} \frac{\phi(x^1 + \epsilon, x^2, \ldots, x^n) - \phi(x^1, x^2, \ldots, x^n)}{\epsilon} \tag{1.14}$$

If all n partial derivatives exist then the function is said to be *differentiable*.

The partial derivatives offer n different ways to differentiate our scalar field. We will sometimes write this as

$$\partial_i \phi = \frac{\partial \phi}{\partial x^i} \tag{1.15}$$

where the ∂_i can be useful shorthand when doing long calculations. While the notation of the partial derivative tells us what's changing it's just as important to remember what's kept fixed. If, at times, there's any ambiguity this is sometimes highlighted by writing

$$\left(\frac{\partial \phi}{\partial x^1}\right)_{x^2, \ldots, x^n}$$

where the subscripts tell us what remains unchanged as we vary x^1. We won't use this notation in these lectures since it should be obvious what variables are being held fixed.

The n different partial derivatives can be packaged together into a vector field. To do this, we introduce the orthonormal basis of vectors $\{\mathbf{e}_i\}$ associated to the coordinates x^i. The *gradient* of a scalar field is then a vector field, defined as

$$\nabla \phi = \frac{\partial \phi}{\partial x^i} \, \mathbf{e}_i \tag{1.16}$$

where we're using the summation convention in which we implicitly sum over the repeated $i = 1, \ldots, n$ index.

Because $\nabla \phi$ is a vector field, it may be more notationally consistent to write it in bold font as $\boldsymbol{\nabla} \phi$. However, I'll stick with $\nabla \phi$. There's no ambiguity here because the symbol ∇ only ever means the gradient, never anything else, and so is *always* a vector. It's one of the few symbols in mathematics and physics whose notational meaning is fixed.

For scalar fields $\phi(x, y, z)$ in $\mathbb{R}^3$, the gradient is

$$\nabla\phi = \frac{\partial\phi}{\partial x}\,\hat{\mathbf{x}} + \frac{\partial\phi}{\partial y}\,\hat{\mathbf{y}} + \frac{\partial\phi}{\partial z}\,\hat{\mathbf{z}}$$

where we've written the orthonormal basis as $\{\mathbf{e}_i\} = \{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$.

There's a useful way to view the vector field $\nabla\phi$. To see this, note that if we want to know how the function ϕ changes in a given direction $\hat{\mathbf{n}}$, with $|\hat{\mathbf{n}}| = 1$, then we just need to take the inner product $\hat{\mathbf{n}} \cdot \nabla\phi$. This is known as the *directional derivative* and sometimes denoted $D_{\mathbf{n}}\phi = \hat{\mathbf{n}} \cdot \nabla\phi$. Obviously the directional derivative is maximal at any point $\mathbf{x}$ when $\hat{\mathbf{n}}$ lies parallel to $\nabla\phi(\mathbf{x})$. But this is telling us something important: at each point in space, the vector $\nabla\phi(\mathbf{x})$ is pointing in the direction in which $\phi(\mathbf{x})$ changes most quickly.

1.3.2 Back to Conservative Fields

First a definition. A vector field $\mathbf{F}$ is called *conservative* if it can be written as

$$\mathbf{F} = \nabla\phi$$

for some scalar field ϕ which, in this context, is referred to as a *potential.* (The odd name "conservative" derives from the conservation of energy in Newtonian mechanics we will see the connection to this below.) Finally, we can answer the question that we introduced at the beginning of this section: when is a line integral independent of the path?

Claim: The line integral around any closed curve vanishes if and only if $\mathbf{F}$ is conservative.

Proof: Consider a conservative vector field of the form $\mathbf{F} = \nabla\phi$. We'll integrate this along a curve C that interpolates from point $\mathbf{a}$ to point $\mathbf{b}$, with parameterisation $\mathbf{x}(t)$. We have

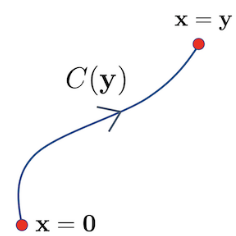
$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_C \nabla\phi \cdot d\mathbf{x} = \int_{t_a}^{t_b} \frac{\partial\phi}{\partial x^i}\,\frac{dx^i}{dt}\,dt = \int_{t_a}^{t_b} \frac{d}{dt}\phi(\mathbf{x}(t))\,dt$$

where the last equality follows from the chain rule. But now we have the integral of a total derivative, so

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \Big[\phi(\mathbf{x}(t))\Big]_{t_a}^{t_b} = \phi(\mathbf{b}) - \phi(\mathbf{a})$$

which depends only on the end points as promised.

Conversely, given the vector field $\mathbf{F}$ whose integral vanishes when taken around any closed curve, it is always possible to construct a potential ϕ. We first choose a value of ϕ at the origin. There's no unique choice here, reflecting the fact that the potential ϕ is only defined up to an overall constant. We can take $\phi(\mathbf{0}) = 0$. Then, at any other point $\mathbf{y}$, we define $\mathbf{x} = \mathbf{y}$



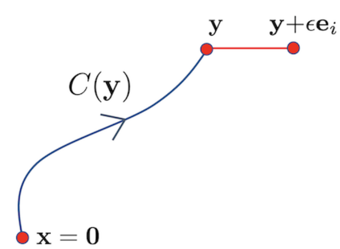
$$\phi(\mathbf{y}) = \int_{C(\mathbf{y})} \mathbf{F} \cdot d\mathbf{x}$$

where $C(\mathbf{y})$ is a curve that starts at the origin and ends at the point $\mathbf{y}$ as shown in the figure above. Importantly, by assumption $\oint \mathbf{F} \cdot d\mathbf{x} = 0$, so it doesn't matter which curve C we take: they all give the same answer.

It remains only to show that $\nabla \phi = \mathbf{F}$. This is straightforward. Reverting to our original definition of the partial derivative (1.14), we have

$$\frac{\partial \phi}{\partial x^i}(\mathbf{y}) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[\int_{C(\mathbf{y}+\epsilon \mathbf{e}_i)} \mathbf{F} \cdot d\mathbf{x} - \int_{C(\mathbf{y})} \mathbf{F} \cdot d\mathbf{x} \right]$$

The first integral goes along $C(\mathbf{y})$, and then continues along the red line shown in the figure to the right. Meanwhile, the second integral goes back along $C(\mathbf{y})$. The upshot is that the difference between them involves only the integral along the red line



$$\frac{\partial \phi}{\partial x^i}(\mathbf{y}) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{\text{red line}} \mathbf{F} \cdot d\mathbf{x}$$

The red line is taken to be the straight line in the x^i direction. This means that the line integral projects onto the F_i component of the vector $\mathbf{F}$. Since we're integrating this over a small segment of length ϵ, the integral gives $\int_{\text{red line}} \mathbf{F} \cdot d\mathbf{x} \approx F_i \epsilon$ and, after taking the limit $\epsilon \to 0$, we have

$$\frac{\partial \phi}{\partial x^i}(\mathbf{y}) = F_i(\mathbf{y})$$

This is our desired result $\nabla \phi = \mathbf{F}$. $\square$

It's clear that the result above is closely related to the fundamental theorem of calculus: the line integral of a conservative vector field is the analog of the integral of a total derivative and so is given by the end points. We'll meet more analogies along the same lines as we proceed.

Given a vector field $\mathbf{F}$, how can we tell if there's a corresponding potential so that we can write $\mathbf{F} = \nabla\phi$? There's one straightforward way to check: for a conservative vector field, the components $\mathbf{F} = F_i \mathbf{e}_i$ are given by

$$F_i = \frac{\partial \phi}{\partial x^i}$$

Differentiating again, we have

$$\frac{\partial F_i}{\partial x^j} = \frac{\partial^2 \phi}{\partial x^i x^j} = \frac{\partial F_j}{\partial x^i} \tag{1.17}$$

where the second equality follows from the fact that the order of partial derivatives doesn't matter (at least for suitably well behaved functions). This means that a necessary condition for $\mathbf{F}$ to be conservative if that $\partial_i F_j = \partial_j F_i$. Later in these lectures we will see that (at least locally) this is actually a sufficient condition.

An Example

Consider the (totally made up) vector field

$$\mathbf{F} = (3x^2 y \sin z, x^3 \sin z, x^3 y \cos z)$$

Is this conservative? We have $\partial_1 F_2 = 3x^2 \sin z = \partial_2 F_1$ and $\partial_1 F_3 = 3x^2 y \cos z = \partial_3 F_1$ and, finally, $\partial_2 F_3 = x^3 \cos z = \partial_3 F_2$. So it passes the derivative test. Indeed, it's not then hard to check that

$$\mathbf{F} = \nabla\phi \quad \text{with} \quad \phi = x^3 y \sin z$$

Knowing this makes it trivial to evaluate the line integral $\int_C \mathbf{F} \cdot d\mathbf{x}$ along any curve C since it is given by $\phi(\mathbf{b}) - \phi(\mathbf{a})$ where $\mathbf{a}$ and $\mathbf{b}$ are the end points of C.

Exact Differentials

There is a slightly different and more abstract way of phrasing the idea of a conservative vector field. First, given a function $\phi(\mathbf{x})$ on $\mathbb{R}^n$, the *differential* is defined to be

$$d\phi = \frac{\partial \phi}{\partial x^i}\, dx^i = \nabla\phi \cdot d\mathbf{x}$$

It's a slightly formal object, obviously closely related to the derivative. The differential is itself a function of $\mathbf{x}$ and captures how much the function ϕ changes as we move in any direction.

Next, consider a vector field $\mathbf{F}(\mathbf{x})$ on $\mathbb{R}^n$. We can take the inner product with an infinitesimal vector to get the object $\mathbf{F} \cdot d\mathbf{x}$. In fancy maths language, this is called a *differential form*. (Strictly it's an object known as a differential one-form) It's best to think of $\mathbf{F} \cdot d\mathbf{x}$ as something that we should integrate along a curve.

A differential form is said to be *exact* if it can be written as

$$\mathbf{F} \cdot d\mathbf{x} = d\phi$$

for some function ϕ. This is just a rewriting of our earlier idea: a differential is exact if and only if the vector field is conservative. In this case, it takes the form $\mathbf{F} = \nabla\phi$ and so the associated differential is

$$\mathbf{F} \cdot d\mathbf{x} = \frac{\partial \phi}{\partial x^i}\, dx^i = d\phi$$

where the last equality follows from the chain rule.

1.3.3 An Application: Work and Potential Energy

There's a useful application of these ideas in Newtonian mechanics. The trajectory $\mathbf{x}(t)$ of a particle is governed by Newton's second law which reads

$$m\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$$

where, in this context, $\mathbf{F}(\mathbf{x})$ can be thought of as a force field. An important concept in Newtonian mechanics is the kinetic energy of a particle, $K = \frac{1}{2}m\dot{\mathbf{x}}^2$. (This is more often denoted as T in theoretical physics.) As the particle's position changes in time, the kinetic energy changes as

$$K(t_2) - K(t_1) = \int_{t_1}^{t_2} \frac{dK}{dt}\, dt = \int_{t_1}^{t_2} m\dot{\mathbf{x}} \cdot \ddot{\mathbf{x}}\, dt = \int_{t_1}^{t_2} \dot{\mathbf{x}} \cdot \mathbf{F}\, dt = \int_C \mathbf{F} \cdot d\mathbf{x}$$

The line integral of the force $\mathbf{F}$ along the trajectory C of the particle is called the *work done*.

Something special happens for conservative forces. These can be written as

$$\mathbf{F} = -\nabla V \tag{1.18}$$

for some choice of V. (Note: the minus sign is just convention.) From the result above, for a conservative force the work done depends only on the end points, not on the path taken. We then have

$$K(t_2) - K(t_1) = \int_C \mathbf{F} \cdot d\mathbf{x} = -V(t_2) + V(t_1) \quad \Rightarrow \quad K(t) + V(t) = \text{constant}$$

We learn that a conservative force, one that can be written as (1.18), has a conserved energy $E = K + V$. Indeed, it's this conservation of energy that lends it's name to the more general idea of a "conservative" vector field. We'll have use of these ideas in the lectures on Dynamics and Relativity.

1.3.4 A Subtlety

Here's a curious example. Consider the vector field on $\mathbb{R}^2$ given by

$$\mathbf{F} = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$$

Is this conservative? If we run our check (1.17), we find

$$\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x} = \frac{y^2 - x^2}{(x^2+y^2)^2}$$

which suggests that this is, indeed, a conservative field. Indeed, you can quickly check that

$$\mathbf{F} = \nabla\phi \quad \text{with} \quad \phi(x,y) = \tan^{-1}\left(\frac{y}{x}\right)$$

(To see this, write $\tan\phi = y/x$ and recall that $\partial(\tan\phi)/\partial x = (\cos\phi)^{-2}\partial\phi/\partial x = (1 + \tan^2\phi)\partial\phi/\partial x$ with a similar expression when you differentiate with respect to y. A little algebra will then convince you that the above is true.)

Let's now integrate $\mathbf{F}$ along a closed curve C that is a circle of radius R surrounding the origin. We take $\mathbf{x}(t) = (R\cos t, R\sin t)$ with $0 \leq t < 2\pi$ and the line integral is

$$\oint_C \mathbf{F}\cdot d\mathbf{x} = \int_0^{2\pi} \mathbf{F}\cdot\frac{d\mathbf{x}}{dt}\,dt = \int_0^{2\pi}\left(-\frac{\sin t}{R}\cdot(-R\sin t) + \frac{\cos t}{R}\cdot R\cos t\right)dt = 2\pi$$

Well, that's annoying! We've just proven that the integral of any conservative vector field around a close curve C necessarily vanishes, and yet one of our first examples seems to show otherwise! What's going on?

The deal is that $\phi(x,y)$ is *not* a well behaved function on $\mathbb{R}^2$. In particular, it's not continuous along the y-axis: as $x \to 0$ the function ϕ approaches either $+\pi/2$ or $-\pi/2$ depending on whether y/x is positive or negative. Implicit in our previous proof was the requirement that we have a continuous function ϕ, well defined everywhere on $\mathbb{R}^2$. Strictly speaking, a conservative field should have $\mathbf{F} = \nabla\phi$ with ϕ continuous.

Relatedly, $\mathbf{F}$ itself isn't defined everywhere on $\mathbb{R}^2$ because it is singular at the origin. Strictly speaking, $\mathbf{F}$ is only defined on the plane $\mathbb{R}^2$ with the point at the origin removed. We write this as $\mathbb{R}^2 - \{0,0\}$, We learn that we should be careful. The line integral of a conservative vector field around a closed curve C is only vanishing if the vector field is well defined everywhere inside C.

Usually pathological examples like this are of interest only to the most self-loathing of pure mathematicians. But not in this case. The subtlety that we've seen above later blossoms into some of the most interesting ideas in both mathematics and physics where it underlies key aspects in the study of topology. In the above example, the space $\mathbb{R}^2 - \{0, 0\}$ has a different topology from $\mathbb{R}^2$ because in the latter case all loops are contractible, while in the former case there are non-contractible loops that circle the origin. It turns out that one can characterise the topology of a space by studying the kinds of functions that live on it. In particular, the functions that satisfy the check (1.17) but cannot be written as $\mathbf{F} = \nabla\phi$ with a continuous ϕ play a particularly important role, as they encode a lot of information about the topology of the underlying space.