Line Integrals, Force Fields and Potentials

April 4, 2016

Overview

These notes cover the following:

- Work done by motion through a force field
- Line integrals
- Conservative force fields and potentials
- Identifying conservative force fields: the curl and Stokes' theorem
- Inverse square laws and Poisson's equation revisited

Work done

You may recall the high school definition of work done on an object (usually a 'particle' with no spatial extent, and with no ability to rotate) as force times distance travelled. At this point, you will probably not be surprised that this definition is not very general. One obvious objection is that the force may change as the particle travels a given distance — so which force to use in the calculation? There is a second issue: Force is a vector, while distance is a scalar, and so is work done. How does 'direction' enter into the problem?

It turns out that the relative orientation of force and *displacement* (as opposed to the scalar 'distance') is what matters. When force and displacement are aligned, work is done *on* the particle, while force and displacement pointing in opposite directions

leads to the particle doing the work. When force and displacement are perpendicular, no work at all is done.[1]

In other words, if a constant force $\mathbf{F}$ acts on an particle, and the particle travels in a straight line to undergo and displacement $\Delta\mathbf{r}$, work done *by that force* on the particle is

$$W = \mathbf{F} \cdot \Delta\mathbf{r}. \tag{1}$$

At this point, the first objection comes into play: what if $\mathbf{F}$ is not constant, or what if the object does not travel along a straight line?

Note 1 *Often you might be interested in the work done by a particular force, say gravity, even when there are other forces such as friction acting; it is therefore not necessary for* $\mathbf{F}$ *to be the total force on the particle.*

By now, the answer to this kind of problem should be clear: if we split the motion of the particle into sufficiently short segments, we can make sure that force is constant and each segment is a straight line, which allows us to apply (1) to that segment. In other words, if the particle travels a short displacement $\delta\mathbf{r}$ while subject to an approximately constant force $\mathbf{F}$, the the work done on the particle during that displacement is

$$\delta W = \mathbf{F} \cdot \delta\mathbf{r}$$

For a longer particle travelling a longer, more complicated trajectory C subject to a changing force, the total work done can then be found by splitting the path into short segments $\delta\mathbf{r}$, computing δW for each, and summing

$$W = \sum \delta W = \sum \mathbf{F} \cdot \delta\mathbf{r}$$

Of course, this is simply another kind of integral, usually called a line integral,

$$W = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

where the subscript C denotes the path followed by the particle.

In practical terms, we are again forced to try to figure out how to reduce a 'formal' integral (a Riemann sum) to an integral we know how to calculate. The easiest way to do this is to parameterize the curve C: assume that the curve can be written in the form

$$\mathbf{r}(t) = X(t)\mathbf{i} + Y(t)\mathbf{j} + Z(t)\mathbf{k} = (X(t), Y(t), Z(t))$$

[1] Consider a spinning bike wheel. In order to keep the rim going along a circular trajectory, forces have to act on the rim to keep parts of it from travelling along a straight line. These forces are provided in part by the spokes of the wheel. Clearly, there is a force, and there is a 'distance travelled'. There is however no gain or loss in the energy contained in the wheel unless there are friction forces. Hence the work done by the spokes must be zero. This is because the forces are at right angles to the rim of the wheel

where X, Y and Z are given functions of a parameter t. For instance, motion around the unit circle in the xy-plane can be written as

$$X(t) = \cos(t), \qquad Y(t) = \sin(t), \qquad Z(t) = 0. \tag{2}$$

Importantly, t *can* stand for time elapsed along the trajectory, but does not have to. In fact, for some curves, t could simply be one of the coordinates. For instance, a parabola in the xy-plane could be written can be written in parameteric form as

$$X(t) = t, \qquad Y(t) = t^2, \qquad Z(t) = 0,$$

which of course is nothing more than $Y = X^2$

Note 2 *We are deliberately using different symbols X, Y and Z to denote the path of the particle: $\mathbf{F}$ may be a function of general position (x, y, z) (for instance, the gravitational force may depend simply on position), and we want to avoid confusion later between that general position, and the particular path followed by a particle.*

The increment $\delta\mathbf{r}$ along the curve C can then be written in terms of the single variable t as

$$\delta\mathbf{r} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\delta t$$

where

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\mathrm{d}X}{\mathrm{d}t}\mathbf{i} + \frac{\mathrm{d}Y}{\mathrm{d}t}\mathbf{j} + \frac{\mathrm{d}Z}{\mathrm{d}t}\mathbf{k} = \left(\frac{\mathrm{d}X}{\mathrm{d}t}.\frac{\mathrm{d}Y}{\mathrm{d}t}, \frac{\mathrm{d}Z}{\mathrm{d}t}\right).$$

The integral can the be written in the form

$$W = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{t_{initial}}^{t_{final}} \mathbf{F} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t$$

where $t = t_{initial}$ and $t = t_{final}$ signify the start and end points on the curve C. If $\mathbf{F}$ is explicitly a function of position (x, y, z), or of position (x, y, z) and, explicitly, of the parameter t (which is quite plausible if t is in fact time, and the force changes over time), then the integral is straightforward to set up. We just have to remember that $\mathbf{F}(x, y, z, t)$ is evaluated at $(X(t), Y(t), Z(t), t)$ and write

$$W = \int_{t_{initial}}^{t_{final}} \mathbf{F}(X(t), Y(t), Z(t), t) \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t \tag{3}$$

Example 1 *Consider the a particle going around the unit circle as described by* (2)*. Consider a force $\mathbf{F}$ being exerted on it that takes the form*

$$\mathbf{F}(x, y, z, t) = -y\mathbf{i} + x\mathbf{j}$$

Compute the work done by $\mathbf{F}$ on the particle in a single revolution.

A single revolution takes us from $t_{initial} = 0$ to $t_{final} = 2\pi$, We also have

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = -\sin(t)\mathbf{i} + \cos(t)\mathbf{j}, \qquad \mathbf{F}(X(t), Y(t), Z(t), t) = -\sin(t)\mathbf{i} + \cos(t)\mathbf{j}$$

and so

$$\mathbf{F} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \sin^2(t) + \cos^2(t) = 1$$

Therefore

$$W = \int_{t_{initial}}^{t_{final}} \mathbf{F} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t = \int_0^{2\pi} 1\,\mathrm{d}t = 2\pi.$$

Exercise 1 *Consider a particle following the path given by*

$$x = \cos(t), \qquad y = \sin(t), \qquad z = t$$

and consider a force $\mathbf{F}$ *applied to the particle in the form*

$$\mathbf{F} = -y\mathbf{i}$$

Sketch the path of the particle. Compute the work done by $\mathbf{F}$ *on the particle between* $t = 0$ *and* $t = 4\pi$

Note 3 *If* $\mathbf{F}$ *is the* total *force acting on the particle, then a link can easily be made between the work done as defined by the line integral and the change in* kinetic *energy of the particle. Let* t *be time, and consider*

$$W = \int_C \mathbf{F} \cdot \,\mathrm{d}\mathbf{r} = \int_{t_{initial}}^{t_{final}} m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t$$

where we have made use of Newton's second law,

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \mathbf{F}$$

where $\mathbf{F}$ *is the total force on the particle, while* m *and* $\mathbf{v}$ *are mass and velocity of the particle. We also know that*

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$$

so that

$$W = \int_{t_{initial}}^{t_{final}} m\mathbf{v} \cdot \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}\,\mathrm{d}t$$

It is straightforward to see that, if $\mathbf{v}(t) = V_x(t)\mathbf{i} + V_y(t)\mathbf{j} + V_z(t)\mathbf{k}$*, that*

$$\mathbf{v} \cdot \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = V_x\frac{\mathrm{d}V_x}{\mathrm{d}t} + V_y\frac{\mathrm{d}V_y}{\mathrm{d}t} + V_z\frac{\mathrm{d}V_z}{\mathrm{d}t} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(V_x^2 + V_y^2 + V_z^2\right) = \frac{1}{2}\frac{\mathrm{d}|\mathbf{v}|^2}{\mathrm{d}t}$$

Hence

$$W = \int_{t_{initial}}^{t_{final}} \frac{1}{2}m\frac{\mathrm{d}|\mathbf{v}|^2}{\mathrm{d}t}\,\mathrm{d}t = \frac{1}{2}m\left|\mathbf{v}\left(t_{final}\right)\right|^2 - \frac{1}{2}m\left|\mathbf{v}\left(t_{initial}\right)\right|^2,$$

where the right-hand side is the change in kinetic energy $m|\mathbf{v}|^2/2$ *between* $t_{initial}$ *and* t_{final}

Curve length and arc length coordinates

Line integrals have uses beyond computing 'work done'. For instance, you might want to compute the length of a curve. Naturally, you would define that length to be

$$L = \sum |\delta \mathbf{r}|,$$

that is, the sum of the lengths of individual segments of the curve C, where

$$|\delta \mathbf{r}| = \sqrt{\delta x^2 + \delta y^2 + \delta z^2}.$$

Again, this integral is fairly straightforward to do if we have parameterized the curve, because we can write

$$|\delta \mathbf{r}| = \left| \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \right| \delta t,$$

assume that δt is positive. Hence

$$L = \int_{t_{initial}}^{t_{final}} \left| \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \right| \mathrm{d}t$$

For reasons that are hopefully obvious, the parameter t is called an *arc length coordinate* if $|\mathbf{r}| = \delta t$, or

$$\left| \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \right| = 1$$

In that case,

$$L = \int_{t_{initial}}^{t_{final}} 1 \, \mathrm{d}t = t_{final} - t_{initial}$$

Arc length coordinates are often denoted by the symbol s rather than t. Also, when parameterized in terms of an arc length coordinate s, the derivative $\mathrm{d}\mathbf{r}/\,\mathrm{d}s$ is simply the unit tangent to the curve C (see also the notes on the gradient operator): It should be clear that the displacement

$$\delta \mathbf{r} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \delta s$$

points in a direction that is tangential to the curve, so that $\mathrm{d}\mathbf{r}/\,\mathrm{d}s$ is tangential to the curve. The requirement that it have unit length then ensures that it is the unit tangent.

Exercise 2 *Compute the length of the curve given by*

$$x = \cos(t), \qquad y = \sin(t), \qquad z = t$$

between $t = 0$ and $t = 2\pi$. Is t an arc length coordinate in this case?

Note 4 *A cautionary tale: in general*

$$\left|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right| \not\equiv \frac{\mathrm{d}|\mathbf{r}|}{\mathrm{d}t}$$

Consider the unit circle example, (2)*, for which*

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = -\sin(t)\mathbf{i} + \cos(t)\mathbf{j}$$

so

$$\left|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right| = 1$$

However, $|\mathbf{r}| = 1$*, so*

$$\frac{\mathrm{d}|\mathbf{r}|}{\mathrm{d}t} = 0.$$

Note that it is possible in principle to transform from any given parameterization $\mathbf{r}(t)$ to a parameterization in terms of an arc length coordinate. Define $s(t)$ as a function of t through[2]

$$s(t) = \int_{t_0}^{t} \left|\frac{\mathrm{d}\mathbf{r}(t')}{\mathrm{d}t'}\right| \mathrm{d}t',$$

where t_0 is fixed but arbitrary. To show s is an arc length coordinate: differentiate in the usual way respect to t on both sides to get, by the fundamental theorem of calculus,

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \left|\frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t}\right|$$

or

$$1 = \left|\frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t}\right| / \frac{\mathrm{d}s}{\mathrm{d}t} = \left|\frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t}\right| \frac{\mathrm{d}t}{\mathrm{d}s}.$$

By $\mathrm{d}t/\,\mathrm{d}s >$, so

$$1 = \left|\frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}s}\right| = \left|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right|$$

by the chain rule, so s is indeed an arc length coordinate.. All that is left to do, at least in principle, is then invert the resulting equation for t as a function of s. This is often not possible analytically, as the next example for a simple parabola shows.

Example 2 *Find an arc length coordinate for the curve*

$$X(t) = t, \qquad Y(t) = \frac{t^2}{2}.$$

[2]Recall that, if we have t in the limit of an integral, we cannot use t as the variable of integration; as a result, we switch to t' as the dummy variable here

Putting $t_0 = 0$, we have $|\,\mathrm{d}\mathbf{r}/\,\mathrm{d}t| = \sqrt{1+t^2}$ and

$$s = \int_0^t \sqrt{1+t'^2}\,\mathrm{d}t'$$

Put $t' = \sinh(u)$, $\mathrm{d}t' = \cosh(u)\,\mathrm{d}u$,

$$\begin{aligned} s(t) &= \int_0^{\sinh^{-1}(t)} \cosh^2(u)\,\mathrm{d}u \\ &= \int_0^{\sinh^{-1}(t)} \frac{1}{2}\left(\cosh(2u)+1\right)\,\mathrm{d}u \\ &= \left[\frac{1}{2}u + \frac{1}{4}\sinh(2u)\right]_0^{\sinh^{-1}(t)} \\ &= \left[\frac{1}{2}u + \frac{1}{2}\sinh(u)\cosh(u)\right]_0^{\sinh^{-1}(t)} \\ &= \frac{1}{2}\sinh^{-1}(t) + \frac{1}{2}t\sqrt{1+t^2} \end{aligned}$$

We see that we can find s in terms of t (or equivalently, in terms of $X = t$). This is the length of the parabola with x-coordinates between 0 and t. Inverting for t is however not possible analytically, on account of the inverse hyperbolic sine.

Force fields, conservative forces, and potentials

The fields we have met so far have been things like temperature, density, mass or heat flux, etc. These all appear naturally in the physics of continua, where mass, energy, momentum and the like are spread out in space. There is another class of fields that are important, especially in geophysics. These are *force fields*, like gravity, electrical and magnetic fields. A particle in such fields experiences a force depending on position and time (which the strength and direction of the field depends on, in general), and on attributes of the particle like mass, charge and velocity. If mass is m, charge is q and velocity is $\mathbf{v}$, the the forces due to the gravitational field $\mathbf{g}$, electrical field $\mathbf{E}$ and magnetic field $\mathbf{B}$ are

$$\mathbf{F}_g = m\mathbf{g}, \qquad \mathbf{F}_e = q\mathbf{E}, \qquad \mathbf{F}_m = q\mathbf{B}\times\mathbf{v}.$$

These fields are important in geophysics in particular because they are also *generated* by masses, charges and currents (as well as acting on masses, charges, and moving charges, respectively). This means that by measuring these force fields, we may be able to learn something about inaccessible masses, charges and currents hidden in the subsurface, where they are difficult to access directly. The processes of using measurements to infer the masses, charges and currents in the subsurface is usually

referred to as 'inverse modelling'. Before we can even begin to understand what that means (and that won't happen in this course), you need to understand the 'forward model': how does a given distribution of mass or charge or current generate a gravitational, electrical or magnetic field. You may recall some very simple formulas for point masses, point charges and possibly even very simple currents. The point of these notes is to lay the groundwork for more general models of how much more complicated distributions of mass, charge and current in space generate the corresponding force fields.

We will focus on the gravitational and electrical fields. This is partly because the forces due to these fields take a simpler form: in both cases, the forces $\mathbf{F}_g$ and $\mathbf{F}_e$ are simply proportional to the gravitational and electrical fields, while the magnetic force also depends on velocity.

Take the simplest case of static force fields, so that $\mathbf{g}$, $\mathbf{E}$ and $\mathbf{B}$ do not depend on time, but only on position. One thing you are hopefully aware of is that the fundamental forces of nature (like gravitational, electrical and magnetic forces) are *conservative*: they cannot create or destroy energy. Instead they can at most transform energy from one form to another. A simple consequence is that, if we take a particle on some 'closed' path C through the force field and return it to its original position with its original velocity, the amount of work done must be zero. Specifically, the amount of work done against the force created by the static force field in traversing the closed path C must be zero:

$$W = \oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0, \tag{4}$$

where the symbol '$\oint$' is frequently used to indicate a line integral is taken over a closed path (or 'closed loop'). In other words, we require that force fields can lead to forces $\mathbf{F}$ that integrate to zero over *any* closed loop as per (4). We would like to have a simple criterion that tells us whether a particular force field satisfies this requirement.

Below, we will focus purely on forces fields like the gravitational and electrical fields, in which $\mathbf{F}$ is given by a scalar property of the particle being moved (mass, charge) times the force field. This is because it is easy to demonstrate that magnetic forces are always conservative, as the next exercise shows.

Exercise 3 *Let the parameter t that we used to define the curve C through the function $\mathbf{r}(t)$ be time. With $\mathbf{F} = q\mathbf{B} \times \mathbf{v}$, show that*

$$W = \int_C \mathbf{F} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t = 0$$

for any *curve C, not just a closed loop. In other words, show that a magnetic force never does any work on a simple charged particle.*

Before we get onto that requirement that ensures (4) holds for all closed loops C, we need to understand the key consequence of that (4): it follows from (4) that the integral

$$\int_{C_{AB}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

along a path C_{AB} from point A to point B depends only on the positions of points A and B, but not on the path. To see this, define C_{BA} be the path C_{AB}, but traversed in the opposite direction (from B to A rather than A to B). Consider two different ways of getting from A to B, labelled $C_{1,AB}$ and $C_{2,AB}$. Then traversing $C_{1,AB}$ followed by $C_{2,BA}$ constitutes a closed loop C, so that

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{C_{1,AB}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \int_{C_{2,BA}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$$

if (4) holds. But, from the definition of the line integral, it is straightforward to see that

$$\int_{C_{2,BA}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = -\int_{C_{2,AB}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

Essentially, in equation (3), reversing the direction in which a path is traversed is equivalent to swapping the limits $t_{initial}$ and t_{final}, which is the same as changing the sign of the integral. But then

$$\int_{C_{1,AB}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} - \int_{C_{2,AB}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$$

or

$$\int_{C_{1,AB}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{C_{2,AB}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

The integral along the two paths, $C_{1,AB}$ and $C_{2.AB}$ is the same and can therefore depend only on the positions of A and B.

This result allows us to define a *potential* associated with the force field $\mathbf{F}$. If we pick a fixed starting point O with position vector $\mathbf{r}_0$, then we can define the potential $\Phi(\mathbf{r}_A)$ at a point A to be the line integral

$$\Phi(\mathbf{r}_A) = -\int_{C_{OA}} \mathbf{F} \cdot \mathrm{d}\mathbf{r},$$

which is the amount of work done *against*[3] the force $\mathbf{F}$ in going from the point O to the point A regardless of the path taken. Defining the point O differently, say moving it from O to O', simply shifts the potential by a constant amount. By picking a path

[3]This is the usual sign convention, as it ensures that the work done is amount of energy stored: carry a rock up a mountain and you are doing work against gravity, which ensures you are storing gravitational potential energy in the rock.

that goes from O' to A via O, we can write the new potential Φ' defined relative O' in terms of the original potential Φ defined relative to O as

$$\Phi'(\mathbf{r}_A) = -\int_{C_{O'A}} \mathbf{F}\cdot\,\mathrm{d}\mathbf{r} = -\int_{C_{O'O}} \mathbf{F}\cdot\,\mathrm{d}\mathbf{r} - \int_{C_{OA}} \mathbf{F}\cdot\,\mathrm{d}\mathbf{r} = -\Phi(\mathbf{r}_{O'}) + \Phi(\mathbf{r}_A),$$

so the two differ only through the term $\int_{C_{O'O}} \mathbf{F}\cdot\,\mathrm{d}\mathbf{r}$, which is determined purely by the shift in reference point from O to O'. For technical reasons, for electrical and gravitational fields, it is common to take the point $\mathbf{r}_O$ to lie at very long distances from all net electrical charges or masses (i.e. 'at infinity'). This is not very relevant at the present stage however.

The potential has a rather important property, which is that we can write $\mathbf{F}$ as its negative gradient:

$$\mathbf{F} = -\nabla\Phi.$$

This is essentially one version of the fundamental theorem of calculus for line integrals. It is straightforward to prove from the fundamental theorem of calculus for single variables and the chain rule. Let $\mathbf{r}(s)$ define a trajectory C through space starting at the point O at $s = 0$, where s is an arc length coordinate. Then, for any point along the trajectory, we have

$$\Phi(\mathbf{r}(s)) = -\int_0^s \mathbf{F}(\mathbf{r}(s'))\cdot\frac{\mathrm{d}\mathbf{r}(s')}{\mathrm{d}s'}\,\mathrm{d}s'.$$

As before, if s features in the limits of integration, we have to use a different dummy variable (here, s') to integrate with respect to. Now differentiate both sides with respect to s. Recall that the left-hand side is nothing more than

$$\Phi(\mathbf{r}(s)) = \Phi(X(s), Y(s), Z(s))$$

The chain rule therefore gives

$$\frac{\mathrm{d}}{\mathrm{d}s}\Phi(\mathbf{r}(s)) = \frac{\partial\Phi}{\partial X}\frac{\mathrm{d}X}{\mathrm{d}s} + \frac{\partial\Phi}{\partial Y}\frac{\mathrm{d}Y}{\mathrm{d}s} + \frac{\partial\Phi}{\partial Z}\frac{\mathrm{d}Z}{\mathrm{d}s} \tag{5}$$

$$= \nabla\Phi\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}. \tag{6}$$

Differentiating the integral on the right, note that we are differentiating an integral of a single variable s' with respect to its upper limit s, which is the second form of the fundamental theorem of calculus:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(x')\,\mathrm{d}x' = f(x).$$

Hence

$$-\frac{\mathrm{d}}{\mathrm{d}s}\int_0^s \mathbf{F}(\mathbf{r}(s'))\cdot\frac{\mathrm{d}\mathbf{r}(s')}{\mathrm{d}s'}\,\mathrm{d}s' = -\mathbf{F}(\mathbf{r}(s))\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s},$$

and therefore

$$\nabla\Phi \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} = -\mathbf{F} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}.$$

Recall that $\mathrm{d}\mathbf{r}/\,\mathrm{d}s = \hat{\mathbf{t}}$ is a unit tangent to the trajectory, in this case at the point $\mathbf{r} = \mathbf{r}(s)$:

$$\nabla\Phi \cdot \hat{\mathbf{t}} = -\mathbf{F} \cdot \hat{\mathbf{t}}.$$

The result however has to hold for *any* trajectory connecting O to the point $\mathbf{r}$. It is clear that I can choose trajectories such that $\hat{\mathbf{t}} = \mathbf{i}$, or $\hat{\mathbf{t}} = \mathbf{j}$ or $\hat{\mathbf{t}} = \mathbf{k}$. In other words,

$$\nabla\Phi \cdot \mathbf{i} = -\mathbf{F} \cdot \mathbf{i}, \qquad \nabla\Phi \cdot \mathbf{j} = -\mathbf{F} \cdot \mathbf{j}, \qquad \nabla\Phi \cdot \mathbf{k} = -\mathbf{F} \cdot \mathbf{k}.$$

The x-components of $\nabla\Phi$ and $-\mathbf{F}$ are equal to each other, and the same is true for the $y-$ and z-components. But that means the two vectors are the same, and so

$$\mathbf{F} = -\nabla\Phi.$$

The converse of this result is much simpler, and is basically the straightforward extension of the fundamental theorem of calculus to line integrals: If $\mathbf{F} = -\nabla\Phi$ and $\Phi(\mathbf{r}_O) = 0$, then it is very easy to show that

$$\Phi(\mathbf{r}_A) = -\int_{C_{OA}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

On the right-hand side, we have, writing in terms of an arc length coordinate starting at O, with s_A indicating the arc length coordinate at position $\mathbf{r}_A$,

$$\begin{aligned}
-\int_{C_{OA}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_{C_{OA}} \nabla\Phi \cdot \mathrm{d}\mathbf{r} \\
&= \int_0^{s_A} -\nabla\Phi \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\,\mathrm{d}s \\
&= \int_0^{s_A} -\left(\frac{\partial\Phi}{\partial X}\frac{\mathrm{d}X}{\mathrm{d}s} + \frac{\partial\Phi}{\partial Y}\frac{\mathrm{d}Y}{\mathrm{d}s} + \frac{\partial\Phi}{\partial Z}\frac{\mathrm{d}Z}{\mathrm{d}s}\right)\mathrm{d}s \\
&= \int_0^{s_A} \frac{\mathrm{d}}{\mathrm{d}s}\Phi(\mathbf{r}(s))\,\mathrm{d}s \\
&= \Phi(\mathbf{r}(s_A)) - \Phi(\mathbf{r}(0)) \\
&= \Phi(\mathbf{r}_A)
\end{aligned}$$

where $\mathbf{r}(s_A) = \mathbf{r}_A$ and $\mathbf{r}(0) = \mathbf{r}_O$ are the locations of point A and O respectively, and $\Phi(\mathbf{r}_O) = 0$ as assumed at the start.[4]

Knowing that a force field is conservative is therefore a powerful tool, as it allows the vector field to be expressed as the gradient of a scalar potential. Instead of needing a model that describes the three components of the vector field, we need only a model for a single scalar field.

[4]Recall also that the first several steps of this derivation were done in the notes on gradients, where we showed that $\nabla \cdot \mathrm{d}\mathbf{r}/\,\mathrm{d}s = \mathrm{d}\Phi/\,\mathrm{d}s$ along a curve C.

The curl and Stokes' theorem

Back to our task of identifying conservative force fields. By a conservative force field, we mean that $\mathbf{F}$ is purely a function of position (and possibly of a scalar property of the moving particle) and that the integral

$$\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$$

around any closed loop C. This turns out to be the case if and only if the so-called *curl* of $\mathbf{F}$ is zero. The curl is the third version of a derivative generated by the vector differential operator ∇. So far we have met the gradient (∇ applied to a scalar field, which gives a vector field because ∇ is vector-valued), and the divergence (∇ applied to a vector field through a 'dot product', which gives a scalar). The curl is natural extension that applies ∇ to a vector field through a 'cross product'. Let

$$\mathbf{F}(x, y, z) = F_x(x, y, z)\mathbf{i} + F_y(x, y, z)\mathbf{j} + F_z(x, y, z)\mathbf{k}.$$

Then the curl of $\mathbf{F}$ is defined as

$$\begin{aligned}\nabla \times \mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} \\ &= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\mathbf{k}\end{aligned}$$

Exercise 4 *Compute the curl of the following vector fields*

1. $\mathbf{F}(x, y, z) = y\mathbf{j} - x\mathbf{i}$
2. $\mathbf{F}(x, y, z) = x\mathbf{i} + y\mathbf{j}$
3. $\mathbf{F}(x, y, z) = x\mathbf{i} - y\mathbf{j}$

Exercise 5 *Compute the sign of each component of the curl (evaluated at the origin) of the vector fields shown in figure 1. Each vector field shown has zero z-component, but note that this does* not *imply that the z-component of the curl is zero. Note that this exercise is similar to figuring out the sign of the divergence of a vector field. Comment on what a 'curl' might mean physically.*

Exercise 6 *Show that*

$$\nabla \times (c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\nabla \times \mathbf{v}_1 + c_2\nabla \times \mathbf{v}_2$$

and

$$\nabla \times \phi\mathbf{v} = \phi\nabla \times \mathbf{v} + (\nabla\phi) \times \mathbf{v}.$$

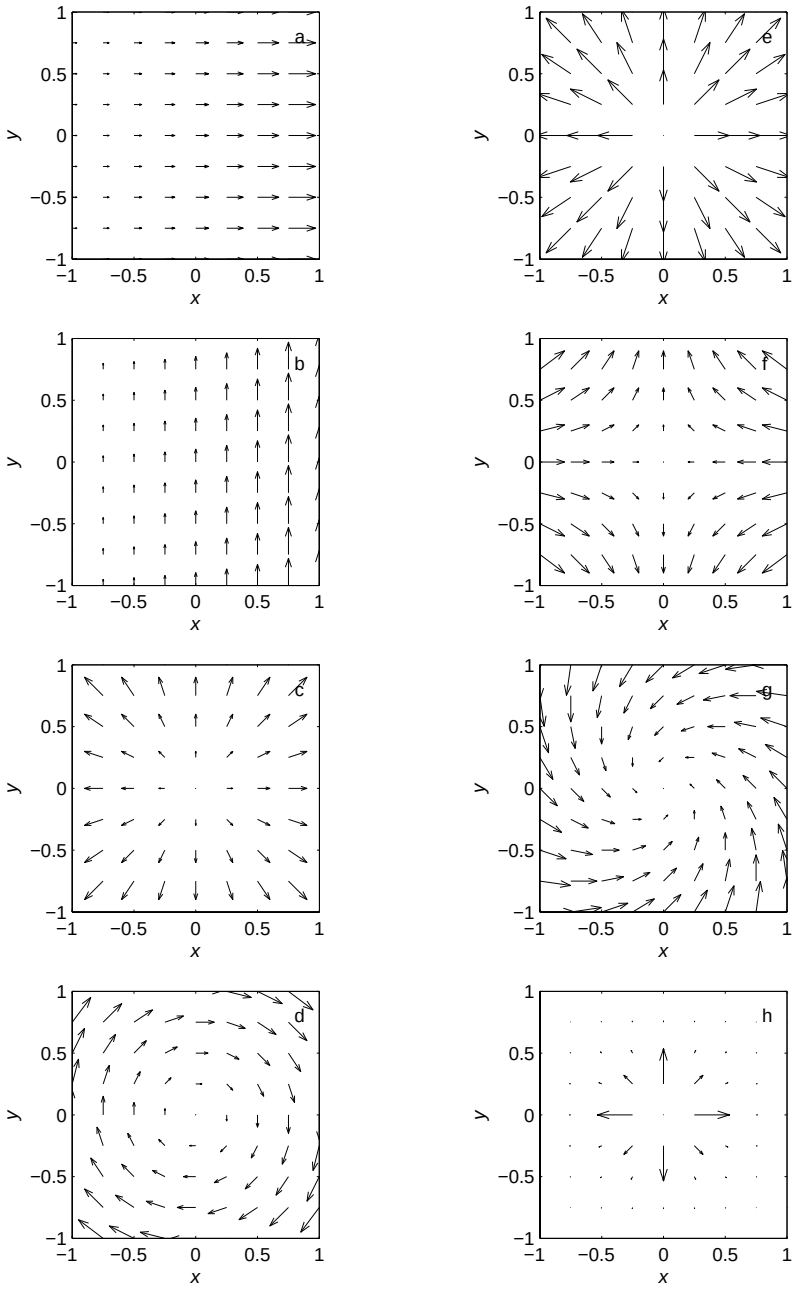

Figure 1: Sample vector fields. Note that each vector field has zero z- ($\mathbf{k}$-) component.

Use the second result to show that $\nabla \times \mathbf{F} = \mathbf{0}$ *if* $\mathbf{F}$ *takes the form*

$$\mathbf{F} = f(r)\hat{\mathbf{r}},$$

where $r = \sqrt{x^2 + y^2 + z^2}$ *and* $\hat{\mathbf{r}} = (x\mathbf{i} + y\mathbf{j} + z\mathbf{k})/r$. *In particular, it follows that an* inverse square law *with* $f(r) = C/r^2$ *has zero curl.*

With this definition, it is easy to prove the 'only if' part of the statement that $\mathbf{F}$ is conservative if and only $\nabla \times \mathbf{F} = \mathbf{0}$. By 'only if', we mean that, if $\mathbf{F}$ being conservative implies that $\nabla \times \mathbf{F} = \mathbf{0}$ (whereas by 'if', we mean that $\nabla \times \mathbf{F} = \mathbf{0}$ implies that $\mathbf{F}$ is conservative).

Given a conservative $\mathbf{F}$, we know from the previous section that there is a potential Φ such that

$$\mathbf{F} = -\nabla\Phi.$$

We can then take the curl of $\mathbf{F}$ to give

$$\nabla \times \mathbf{F} = -\nabla \times \nabla\Phi.$$

You probably recall from ordinary vector algebra that the cross product of a vector with itself is zero. The same turns out to apply to the operator ∇, meaning

$$\nabla \times \mathbf{F} = -\nabla \times \nabla\Phi = \mathbf{0}.$$

To show this, substitute $\mathbf{F} = -\nabla\Phi$ into the definition of the curl above, in the form

$$F_x = -\frac{\partial\Phi}{\partial x}, \qquad F_y = -\frac{\partial\Phi}{\partial y}, \qquad F_z = -\frac{\partial\Phi}{\partial z},$$

so that

$$\begin{aligned}\nabla \times \mathbf{F} &= \left(-\frac{\partial^2\Phi}{\partial y\partial z} + \frac{\partial^2\Phi}{\partial z\partial y}\right)\mathbf{i} + \left(\frac{\partial^2\Phi}{\partial z\partial x} - \frac{\partial^2\Phi}{\partial x\partial z}\right)\mathbf{j} + \left(\frac{\partial^2\Phi}{\partial x\partial y} - \frac{\partial^2\Phi}{\partial y\partial x}\right)\mathbf{k} \\ &= \mathbf{0}\end{aligned}$$

as we can exchange the order of differentiation in each case, so that each component of $\nabla \times \mathbf{F}$ vanishes.

The 'if' part of the statement is much harder to demonstrate. It relies on the last 'big' result in this course, Stokes' theorem. Stokes theorem is to curls what the divergence theorem is to divergences.[5]

Stokes' theorem states the following: Let S be an *open* surface.[6] This means S has a boundary curve C, which obviously takes the form of a closed loop, and let $\hat{\mathbf{n}}$

[5]For completeness of historical digression and naming symmetry, the divergence theorem also sometimes gets named after a dead mathematician: Stokes' theorem is named for George Gabriel Stokes, while the divergence theorem is sometimes referred to as Gauss' theorem, after Carl Friedrich Gauss.

[6]as opposed to the *closed* surface in the divergence theorem

be a unit normal to the surface S. There are two choices for $\hat{\mathbf{n}}$ ('up' and 'down'), but no inward- or outward-pointing direction because S is open and does not enclose a volume. Similarly, there are two orientations in which C can be traversed. Let the orientation in which C is traversed and the orientation of $\hat{\mathbf{n}}$ be linked by a 'right-hand rule': imagine the 'thumb' of the right hand point along C in the direction of traverse and the index finger point into the surface S. Then the middle finger points in the direction of $\hat{\mathbf{n}}$.[7] Stokes theorem then states that

$$\int_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \oint_C \mathbf{F} \cdot \, \mathrm{d}\mathbf{r}.$$

Assuming this result to be true, it is then straightforward to see that $\nabla \times \mathbf{F} = \mathbf{0}$ implies

$$\int_C \mathbf{F} \cdot \, \mathrm{d}\mathbf{r} = \int_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = 0$$

for any closed loop, proving the 'if' part of our statement about conservative vector fields.

Note 5 *Note that Stokes theorem is* consistent *with the divergence theorem. If we turn S into a closed surface, then the bounding curve C has to shrink to zero, meaning $\oint_C \mathbf{F} \cdot \, \mathrm{d}\mathbf{r} = 0$. Then*

$$\int_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = 0.$$

But S is now a a closed surface, so we can apply the divergence theorem and have to arrive at

$$0 = \int_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \int_V \nabla \cdot (\nabla \times \mathbf{F}) \, \mathrm{d}V.$$

It turns out that this guaranteed, because

$$\nabla \cdot \nabla \times \mathbf{F} = 0$$

identically. To show this last identity, use the definition of divergence and curl to find

$$\begin{aligned}
\nabla \cdot \nabla \times \mathbf{F} &= \frac{\partial}{\partial x}\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) \\
&= \frac{\partial^2 F_z}{\partial x \partial y} - \frac{\partial^2 F_y}{\partial x \partial z} + \frac{\partial^2 F_x}{\partial y \partial z} - \frac{\partial^2 F_z}{\partial y \partial x} + \frac{\partial^2 F_y}{\partial z \partial x} - \frac{\partial^2 F_x}{\partial z \partial y} \\
&= 0
\end{aligned}$$

because, again, we can exchange the order of differentiation.

[7] A simpler way to think of this is: imagine S is the interior of the unit circle in the xy-plane, and that the circle itself is traversed in the anti-clockwise direction. The $\hat{\mathbf{n}}$ points in the positive z-direction.

Next, we will prove a limited version of Stokes' theorem. As in our notes on surface integrals, we consider a surface S given by $z = h(x, y)$, whose projection onto the xy-plane is bounded on the left by $x = x_{min}$, on the right by $x = x_{max}$, by $y = y_{max}(x)$ above and by $y = y_{min}(x)$ below (figure 2. This implies that we break the boundary curve C into two parts, one of which we can parameterized in terms of x (rather than t!) through

$$y = y_{min}(x), \qquad z = h(x.y_{min}(x)), \tag{7}$$

and the other through

$$y = y_{max}(x), \qquad z = h(x, y_{max}(x)). \tag{8}$$

Assume that the boundary curve C is traverse in an anticlockwise orientation, It then follows that the first, 'lower' segment of C given by (7) is traversed with x increasing, and the 'upper' segment given by (8) is traversed with x decreasing. With that orientation of C, the unit normal to S is upward-pointing, given by

$$\hat{\mathbf{n}} = \frac{\mathbf{k} - \mathbf{i}\frac{\partial h}{\partial x} - \mathbf{j}\frac{\partial h}{\partial y}}{\sqrt{1 + \left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}.$$

We consider also a force field $\mathbf{F}$ of the form $\mathbf{F} = F_x(x, y, z)\mathbf{i}$ only. Putting $\mathbf{r}_{min}(x) = (x, y_{min}(x), h(x, y_{min}(x))$ and $\mathbf{r}_{max}(x) = (x, y_{max}(x), h(x, y_{max}(x))$, the line integral then becomes

$$\begin{aligned}\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_{x_{min}}^{x_{max}} F_x(\mathbf{r}_{min}(x))\mathbf{i} \cdot \frac{\mathrm{d}\mathbf{r}_{min}}{\mathrm{d}x}\,\mathrm{d}x + \int_{x_{max}}^{x_{min}} F_x(\mathbf{r}_{max}(x))\mathbf{i} \cdot \frac{\mathrm{d}\mathbf{r}_{max}}{\mathrm{d}x}\,\mathrm{d}x \\ &= \int_{x_{min}}^{x_{max}} F_x(x, y_{min}(x), h(x, y_{min}(x)))\,\mathrm{d}x + \int_{x_{max}}^{x_{min}} F_x(x, y_{max}(x), h(x, y_{max}(x)))\,\mathrm{d}x\end{aligned}$$

We also have the curl of $\mathbf{F} = F_x\mathbf{i}$ given by

$$\nabla \times \mathbf{F} = \frac{\partial F_x}{\partial z}\mathbf{j} - \frac{\partial F_x}{\partial y}\mathbf{k}.$$

Use this to compute the surface integral in Stokes' theorem:

$$\begin{aligned}\int (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}}\,\mathrm{d}S &= \int_{x_{min}}^{x_{max}} \int_{y_{min}(x)}^{y_{max}(x)} \left(\left.\frac{\partial F_x}{\partial z}\right|_{z=h(x,y)} \mathbf{j} - \left.\frac{\partial F_x}{\partial y}\right|_{z=h(x,y)} \mathbf{k} \right) \cdot \left(\mathbf{k} - \mathbf{i}\frac{\partial h}{\partial x} - \mathbf{j}\frac{\partial h}{\partial y} \right) \mathrm{d}y\,\mathrm{d}x \\ &= \int_{x_{min}}^{x_{max}} \int_{y_{min}(x)}^{y_{max}(x)} - \left[\left.\frac{\partial F_x}{\partial y}\right|_{z=h(x,y)} + \left.\frac{\partial F_x}{\partial z}\right|_{z=h(x,y)} \frac{\partial h}{\partial y} \right] \mathrm{d}y\,\mathrm{d}x.\end{aligned}$$

Define a new function of x and y only, given by $f(x, y) = F_x(x, y, z = h(x, y))$. Differentiating f with respect to y, we get by the multivariable chain rule that

$$\frac{\partial f}{\partial y} = \frac{\partial F_x}{\partial y} + \frac{\partial F_x}{\partial z}\frac{\partial h}{\partial y}$$

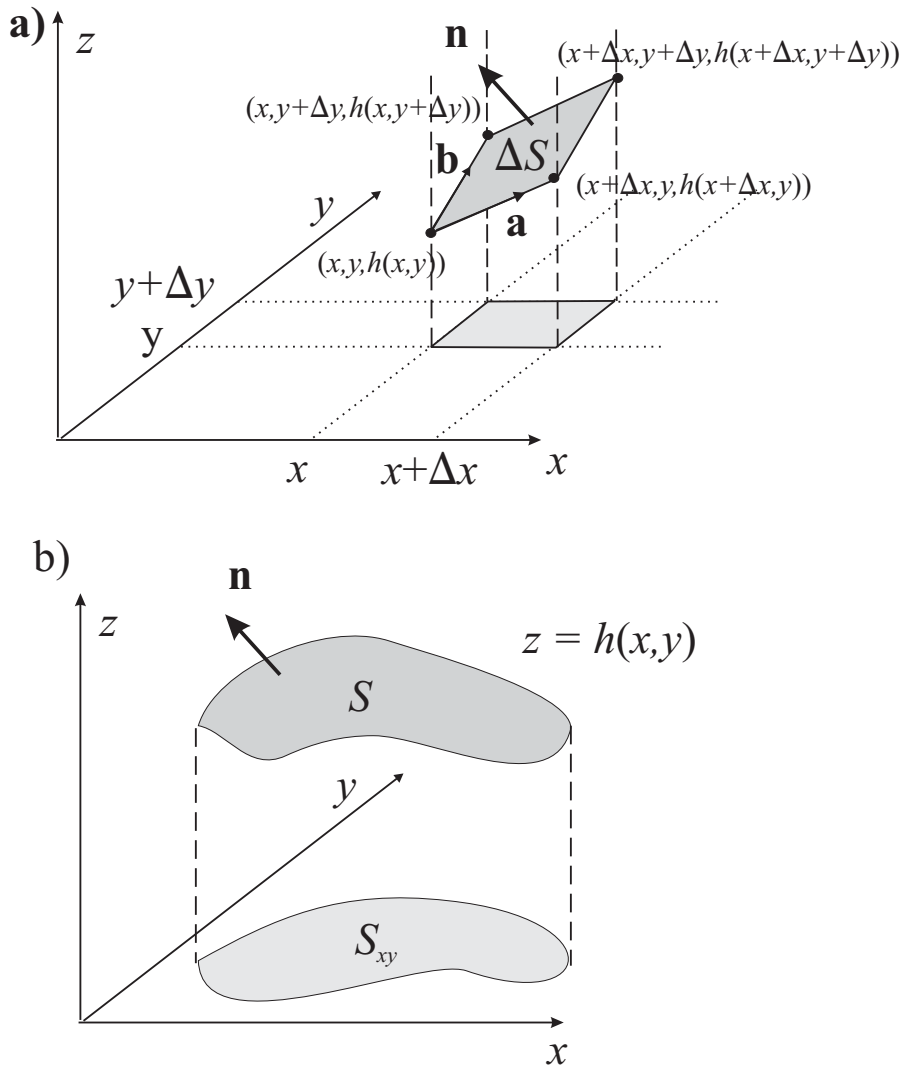

Figure 2: The lower panel shows the geometry we have in mind for computing the line integral $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ (anticlockwise around the perimeter of the surface) and surface integral $\int_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S$ (with the unit normal pointing up).

Hence, using the fundamental theorem of calculus to integrate with respect to y,

$$\begin{aligned}
\int (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S &= \int_{x_{min}}^{x_{max}} \int_{y_{min}(x)}^{y_{max}(x)} -\frac{\partial f}{\partial y} \, \mathrm{d}y \, \mathrm{d}x \\
&= \int_{x_{min}}^{x_{max}} -\left[f(x,y) \right]_{y=y_{min}(x)}^{y=y_{max}(x)} \, \mathrm{d}x \\
&= -\int_{x_{min}}^{x_{max}} f(x, y_{max}(x)) \, \mathrm{d}x + \int_{x_{min}}^{x_{max}} f(x, y_{min}(x)) \, \mathrm{d}x \\
&= \int_{x_{max}}^{x_{min}} f(x, y_{max}(x)) + \int_{x_{min}}^{x_{max}} f(x, y_{min}(x)) \, \mathrm{d}x \\
&= \int_{x_{max}}^{x_{min}} F_x(x, y_{max}(x), h(x, y_{max}(x)) \, \mathrm{d}x + \int_{x_{min}}^{x_{max}} F_x(x, y_{min}(x), h(x, y_{min}(x)) \, \mathrm{d}x
\end{aligned}$$

where we have used the definition of $f(x, y)$ in the last equality. But the last line is exactly the same as the line integral computed previously, so we have

$$\int_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \oint_C \mathbf{F} \cdot \, \mathrm{d}\mathbf{r} \tag{9}$$

if $\mathbf{F} = F_x \mathbf{i}$. In other words, what we have really shown is that

$$\int_S (\nabla \times \mathbf{F}_x \mathbf{i}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \oint_C F_x \mathbf{i} \cdot \, \mathrm{d}\mathbf{r}$$

Note that the derivation of this result has relied on the fundamental theorem of calculus; like the divergence theorem, Stokes' theorem is a vector calculus version of the fundamental theorem of calculus.

The idea is then that there is nothing special about the x-direction, and we should be able to show the same result for $\mathbf{F} = F_y \mathbf{j}$ or $\mathbf{F} = F_z \mathbf{k}$, and summing the results for all three we get

$$\int_S (\nabla \times \mathbf{F}_x \mathbf{i}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S + \int_S (\nabla \times \mathbf{F}_y \mathbf{j}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S + \int_S (\nabla \times \mathbf{F}_z \mathbf{k}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \int_C F_x \mathbf{i} \cdot \mathrm{d}\mathbf{r} + \int_C F_y \mathbf{j} \cdot \mathrm{d}\mathbf{r} + \int_C F_z \mathbf{k} \cdot \mathrm{d}\mathbf{r}.$$

But the integrals on the left-hand and right-hand sides can be combined to give

$$\int_S [\nabla \times (\mathbf{F}_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k})] \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \oint_C (F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k}) \cdot \, \mathrm{d}\mathbf{r} \tag{10}$$

or more simply

$$\int_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \oint_C \mathbf{F} \cdot \, \mathrm{d}\mathbf{r}$$

for arbitrary vector fields $\mathbf{F}$. To really make this work, we do have to get rid of the restriction to very simple surfaces and bounding curves however (as that restriction

made our derivation of Stokes theorem for $\mathbf{F} = F_x\mathbf{i}$ possible in the relatively simple form given above). In general, this can involve both, the surface S or the bounding curve C folding back on themselves. In the following note, we only consider the case of the surface S folding back on itself.

Note 6 *To generalize the result* (9) *to arbitrary surfaces (even for a restrictive vector field like* $\mathbf{F} = F_x\mathbf{i}$*), note first that we can split more complicated surfaces that fold back on themselves (i.e., for which we cannot write* $z = h(x, y)$ *with a single-valued function* h*) into multiple simple ones that can be written in the form* $z = h(x, y)$*. (This is analogous to Note 4 in the notes on the divergence theorem). Consider a surface* S *that folds back on itself once, so it needs to be split into two parts* S_1 *and* S_2*. Then*

$$\int_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \int_{S_1} (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S + \int_{S_2} (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S$$

In terms of the integrals over bounding curves C*,* C_1 *and* C_2*, we now have to account for the part line integrals along* C_1 *and* C_2 *that are not part of the original bounding curve* C*, but that form part of the cut that divides the original surface* S*. Let that cut be* $C_{1,cut}$ *when traversed in the positive direction around* C_1*, and* $C_{2,cut}$ *when traversed in the positive direction around* C_2*. We have*

$$\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \oint_{C_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} - \int_{C_{1,cut}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \oint_{C_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} - \int_{C_{2,cut}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

Note however that $C_{1,cut}$ *and* $C_{2,cut}$ *are the same curve, only traversed in opposite directions. Because the line integral changes sign when the curve is traversed in the opposite direction, we therefore have*

$$\int_{C_{1,cut}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = -\int_{C_{2,cut}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

and the two integrals along the cut cancel. As a result,

$$\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \oint_{C_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \oint_{C_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

As a result, if we can prove Stokes' theorem for the surfaces S_1 *and* S_2 *and there respective bounding curves* C_1 *and* C_2*, it follows that the theorem also holds for the combined surface* S *and its bounding curve* C*.*

The same argument can also be used if the bounding curve C *folds back on itself, so there is not a single* $y = y_{min}(x)$ *or a single* $y = y_{max}(x)$*: in this case, we can cut* S *into two pieces such that each has a single 'lower' boundary* $y_{min}(x)$ *and single 'upper' boundary* $y_{max}(x)$*. If Stokes' theorem holds for each piece, it will also hold for the combined piece. In this way we can reduce the surface* S *we start with to a*

collection of simple surfaces, each of the form for which we proved Stokes' theorem for $\mathbf{F} = F_x\mathbf{i}$. *This allows us to derive to derive* (9) *for arbitrary surfaces with* $\mathbf{F} = F_x\mathbf{i}$, *and we can repeat the procedure analogously (switching the roles of* x *with* y *and* z*) to show that* (9) *holds for arbitrary surfaces with* $\mathbf{F} = F_y\mathbf{j}$ *and* $\mathbf{F} = F_z\mathbf{k}$. *The full version of Stokes theorem then follows, again by adding.*

Inverse square laws and Poisson's equation for electrostatic and gravitational fields

Knowing that a vector field is conservative is not enough to specify the actual form of the vector field. It tells us one simplification — the vector field is the gradient of a scalar field — but goes no further. There is in fact an analogy with the case of divergence: imagine you are told that a flux has zero divergence. That by itself is not enough to tell you what the flux field is. In the case of heat fluxes, we needed an additional constitutive law like Fourier's law to allow us to compute temperature and therefore heat flux.

Here we focus on electrostatic and gravitational fields, and the additional 'constitutive law' that is needed. (Though that phrase is generally not used in this context: 'constitutive laws' are generally empirical and material-dependent, whereas electrical and gravitational fields are fundamental force fields of nature.) In short, our aim is to figure out how electrostatic and gravitational fields are governed by the distribution of charges and masses.

You are likely to recall two facts about electrostatic and gravitational fields from high school. First, the field due to a point charge or a point mass follows an 'inverse square law': the field is proportional to the inverse square of the distance between the point charge or mass and the point at which you want to know the strength of the field. The field also points either directly towards or away from the point charge or mass ('away' being possible only for electrostatic fields, in the case of a positive charge generating the field). The inverse square law for gravitational fields is generally known as Newton's law of gravitation, while its electrostatic counterpart is Coulomb's law.

Second, the effect of multiple charges or masses is additive. For a given point, you simply work out the contribution to the field due to each surrounding charge or mass as though the other charges or masses were absent, and then add all those contributions.

Let us put this in concrete mathematical form for an electrostatic field. Let there be a charge q_i at a location $\mathbf{r}_i = (x_i, y_i, z_i)$, and let $\mathbf{r} = (x, y, z)$ be the location at which you want to know the electrostatic field. The distance between the two points is then

$$|\mathbf{r} - \mathbf{r}_i|$$

and the *strength* of the field at $\mathbf{r}$ due to the charge q_i is inversely proportional to the

square of that distance, and proportional to the size of the charge

$$|\mathbf{E}_i| \propto \frac{q_i}{|\mathbf{r}-\mathbf{r}_i|^2}.$$

We have put a subscript 'i' on $\mathbf{E}_i$ in order to make it easier to add fields later. The constant of proportionality is usually denoted by $1/(4\pi\epsilon_0)$, where ϵ_0 is the 'permittivity of free space', so

$$|\mathbf{E}_i| = \frac{q_i}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}-\mathbf{r}_i|^2}.$$

The constant of proportionality may seem needlessly complicated (why the 'one over'? why 4π?) but its choice will make a lot more sense later. For now, just remember the first fraction is a constant.

The formula above gives us the field *strength*, but the field also has a direction. In the case of an electrostatic field, it points *away* from the charge location $\mathbf{r}_i$ if the charge q_i is positive — charges of the same sign repel each other. The *direction* of the field is in the direction of the vector linking the charge location $\mathbf{r}_i$ to the point $\mathbf{r}$. In other words, its direction is the direction of

$$\mathbf{r}-\mathbf{r}_i.$$

We can construct the vector $\mathbf{E}_i$ as the product of the strength ($|\mathbf{E}_i$—) and the *unit* vector in the direction of the field. That unit vector is obviously

$$\frac{\mathbf{r}-\mathbf{r}_i}{|\mathbf{r}-\mathbf{r}_i|},$$

so that

$$\begin{aligned}\mathbf{E}_i &= \frac{q_i}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}-\mathbf{r}_i|^2}\frac{\mathbf{r}-\mathbf{r}_i}{|\mathbf{r}-\mathbf{r}_i|}\\ &= \frac{q_i}{4\pi\epsilon_0}\frac{\mathbf{r}-\mathbf{r}_i}{|\mathbf{r}-\mathbf{r}_i|^3}.\end{aligned}$$

If there are multiple point charges surrounding the point $\mathbf{r}$, we simply add the effect of each charge. This is why we have labelled $\mathbf{E}_i$, q_i and $\mathbf{r}_i$ above with subscripts 'i'. For a number of charges, we can give each a label $i = 1, 2, \ldots$, and add to get the total field $\mathbf{E}(\mathbf{r})$ as a function of position $\mathbf{r}$ through

$$\mathbf{E}(\mathbf{r}) = \sum_i \frac{q_i}{4\pi\epsilon_0}\frac{\mathbf{r}-\mathbf{r}_i}{|\mathbf{r}-\mathbf{r}_i|^3}. \tag{11}$$

For the equivalent gravitational field with point masses m_i at locations $\mathbf{r}_i$, we would have

$$\mathbf{g}(\mathbf{r}) = -\sum_i Gm_i\frac{\mathbf{r}-\mathbf{r}_i}{|\mathbf{r}-\mathbf{r}_i|^3}. \tag{12}$$

where G is the gravitational constant, and the minus sign takes care of the fact that masses attract rather than repel each other.

Exercise 7 *To show that* $\mathbf{E}$ *and* $\mathbf{g}$ *as stated in* (11) *and* (12) *are both conservative fields, we need to show that they have zero curl. This is an application of the product rule in exercise 6. Let* $r_i(x,y,z) = |\mathbf{r}-\mathbf{r}_i| = \sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}$. *Show that*

$$\nabla r_i = \frac{\mathbf{r}-\mathbf{r}_i}{r_i},$$

and that

$$\nabla \times (\mathbf{r}-\mathbf{r}_i) = \mathbf{0}.$$

Also using the product rule for gradients, $\nabla f(r_i) = f'(r_i)\nabla r_i$, *show that*

$$\nabla \times \frac{\mathbf{r}-\mathbf{r}_i}{|\mathbf{r}-\mathbf{r}_i|^3} = \mathbf{0}$$

and hence that $\nabla \times \mathbf{g} = \nabla \times \mathbf{E} = \mathbf{0}$.

Equations (11) and (12) work well for *point charges*, so the challenge is to generalize the results for $\mathbf{E}$ and $\mathbf{g}$ above to the case of spatially distributed charges or masses — meaning, rather than having discrete point charges and point masses, a finite charge or mass is only contained in a finite volume, the concentration of mass or charge being described by densities as described in the notes on volume integrals.

To understand that generalization, take a second look at the end of the notes on the heat equation. Recall that steady heat conduction associated with a number of discrete point sources of heat gives rise to a heat flux field that also follows an inverse square law. In fact, the formula for the flux is completely analogous to the formulas for $\mathbf{E}$ and $\mathbf{g}$ generated by discrete point charges and masses above. With a constant thermal conductivity k and heat sources of strength Q_i located at $\mathbf{r}_i$ $(i = 1, 2, \ldots)$, the conductive heat flux $\mathbf{q}(\mathbf{r})$ at a location $\mathbf{r}$ is given by

$$\mathbf{q}(\mathbf{r}) = \sum_i \frac{Q_i}{4\pi} \frac{\mathbf{r}-\mathbf{r}_i}{|\mathbf{r}-\mathbf{r}_i|^3} \tag{13}$$

Moreover, in this case the flux $\mathbf{q}$ is generated by the gradient of a 'potential' kT (T being temperature)

$$\mathbf{q} = -k\nabla T = -\nabla(kT), \tag{14}$$

where

$$kT = \sum_i \frac{Q_i}{4\pi} \frac{1}{|\mathbf{r}-\mathbf{r}_i|} + kT_\infty$$

This implies that, like $\mathbf{E}$ and $\mathbf{g}$, the flux $\mathbf{q}$ is also conservative, with $\nabla \times \mathbf{q} = \mathbf{0}$.

The mathematical analogy is especially close with (11). Taking the constant ϵ_0 to the left-hand side of (11), we see that replacing q_i by Q_i and $\epsilon\mathbf{E}$ by $\mathbf{q}$ gives us (13), while we also expect $\epsilon_0\mathbf{E}$ to be the gradient of a potential. In fact, this would be $\epsilon_0\phi$, where ϕ is the ordinary electrostatic potential, so replacing $\epsilon_0\phi$ by kT as well turns

$$\mathbf{E} = -\nabla\phi$$

into

$$\mathbf{q} = -\nabla(kT).$$

Now, to complete the generalization, we know that a point source of heat Q_i (with dimensions of energy per time) is simply an idealization in the world of heat transport. In reality, we expect heat to be produced in a spatially distributed way, described by the 'heat production rate density' a (with dimensions of energy per time and volume). Effectively, Q_i is a high production rate density a concentrated in a small volume δV, with $Q_i = a\delta V$. In the more general case of a spatially distributed heat source, we had the flux determined by conservation of energy through

$$\nabla \cdot \mathbf{q} = a, \qquad \mathbf{q} = -\nabla(kT). \tag{15}$$

Let charge density be denoted by ρ_c (to distinguish it from mass density ρ). The generalization of a point charge is the that q_i is really a highly concentrated charge density ρ_c in a small volume δV, meaning ρ_c is analogous to a above. In short, generalizing from point charges to spatially distributed charges, we expect $\epsilon_0\mathbf{E}$ to satisfy

$$\nabla \cdot \epsilon_0\mathbf{E} = \rho_c, \qquad \epsilon_0\mathbf{E} = -\nabla(\epsilon_0\phi).$$

More succinctly, we get

$$\epsilon_0\nabla \cdot \mathbf{E} = \rho_c, \qquad \mathbf{E} = -\nabla\phi, \tag{16}$$

which put together gives

$$-\epsilon_0\nabla^2\phi = \rho_c.$$

Once more, we have Poisson's equation. Notice how only ϵ_0 appears on the left; this explains the somewhat complicated choice of the constant of proportionality in Coulomb's law.

The fact that a similar generalization for the gravitational field is possible should be clear from the fact that Coulomb's law and Newton's law of gravitation are both inverse square laws, only with different constants. The equivalent to (16) is

$$\nabla \cdot \mathbf{g} = -4\pi G\rho, \qquad \mathbf{g} = -\nabla\phi$$

were ϕ is now the gravitational potential, and ρ the ordinary mass density. Clearly, the gravitational potential ϕ also satisfies Poisson's equation,

$$\nabla^2\phi = 4\pi G\rho.$$

This is sometimes known as Gauss' law of gravitation.

Exercise 8 *Repeat the derivation of* (16) *for gravitational fields. What constants do you need to multiply* $\mathbf{g}$ *and* m_i *by in order to make them equivalent to* $\mathbf{q}$ *and* Q_i *in* (13)*?*

Note 7 *Note that, instead of writing* (16)*, one would often write the equivalent*

$$\epsilon_0 \nabla \cdot \mathbf{E} = \rho_c, \qquad \nabla \times \mathbf{E} = \mathbf{0}, \tag{17}$$

where the second equation implies the existence of a potential. This is because, when the electrical field is not steady — so we no longer have an electrostatic *field — the curl of* $\mathbf{E}$ *is no longer zero but is related to the rate of change of the magnetic field with respect to time. This is the basis of such effects as magnetic induction, and the propagation of electromagnetic waves. The relevant extension of* (17) *leads to* Maxwell's equations*. See a course in electromagnetism — for which you should now be well-prepared.*