Lecture Notes for

MTH5102: CALCULUS III

Dr. Will Sutherland

These notes are updated from previous versions by Malcolm McCallum and M.J. Thompson, with figures by C.D. Murray, and by P. Saha.

School of Mathematical Sciences,
Queen Mary, University of London

September 2010

> "It is one of the most unnatural features of science that the abstract language of mathematics should provide such a powerful tool for describing the behaviour of systems both inanimate, as in physics, and living, as in biology. Why the world should conform to mathematical descriptions is a deep question. Whatever the answer, it is astonishing."
> Lewis Wolpert (1992)

Chapter 1

Introductory material

Last revised: 29 Sep 2010

This chapter gives a quick review of the key parts of the prerequisite courses (Calculus I and II, and Geometry I) which we will actually use in Calculus III, adding some extra material. Those parts which are revision will be without examples.

1.1 Trigonometric functions

1.1.1 Values

(See Thomas 1.6)
We can quickly obtain the value of a trigonometric function for any argument in terms of values for $x \in [0, \frac{1}{2}\pi]$ by remembering a few things. First we have the table

	0°	$30^\circ = \frac{\pi}{6}$ radians	$45^\circ = \frac{\pi}{4}$ rad.	$60^\circ = \frac{\pi}{3}$ rad.	$90^\circ = \frac{\pi}{2}$ rad.
cos	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
sin	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1

To get the sign for other values we can use the mnemonic table

Radians	Degrees	sin	cos	tan	Positive functions
$(0, \frac{1}{2}\pi)$	$(0^\circ, 90^\circ)$	+	+	+	All
$(\frac{1}{2}\pi, \pi)$	$(90^\circ, 180^\circ)$	+	–	–	Sin
$(\pi, \frac{3}{2}\pi)$	$(180^\circ, 270^\circ)$	–	–	+	Tan
$(\frac{3}{2}\pi, 2\pi)$	$(270^\circ, 360^\circ)$	–	+	–	Cos

sometimes called the 'Add Sugar To Coffee' rule – or use Thomas' variant "All Students Take Calculus". (Note: to be entirely accurate we should have special rows in this table for the values $\frac{1}{2}\pi$ etc because at those points one or more of the functions will be zero or unbounded.)

Then we remember what happens when we replace x by $-x$, $x+\pi/2$ or $x+\pi$:

$$\begin{aligned} \cos(-x) &= \cos x, \quad \sin(-x) = -\sin x, \\ \cos(x+\frac{\pi}{2}) &= -\sin x, \quad \sin(x+\frac{\pi}{2}) = \cos x, \\ \cos(x+\pi) &= -\cos x, \quad \sin(x+\pi) = -\sin x. \end{aligned} \tag{1.1}$$

These are very easy to derive from $e^{ix} = \cos x + i\sin x$, remembering that $e^{i\pi/2} = i$, $e^{i\pi} = -1$. Using them in combination we can get

$$\cos(\pi - x) = -\cos x, \quad \sin(\pi - x) = \sin x.$$

and so on.

More generally

$$\cos(x+(m+\frac{1}{2})\pi) = (-1)^{(m+1)}\sin x, \quad \sin(x+(m+\frac{1}{2})\pi) = (-1)^m \cos x, \tag{1.2}$$

$$\cos(x+n\pi) = (-1)^n \cos x, \quad \sin(x+n\pi) = (-1)^n \sin x. \tag{1.3}$$

where m and n are integers. These identities enable us to relate the value we want to a value in the first quadrant (i.e. the range $[0, \frac{1}{2}\pi]$). Remember the special cases for $x = 0$,

$$\cos(n\pi) = (-1)^n, \qquad \sin((n+\frac{1}{2})\pi) = (-1)^n. \tag{1.4}$$

$$\cos((n+1/2)\pi) = 0, \qquad \sin(n\pi) = 0. \tag{1.5}$$

which will turn up regularly later on.

1.1.2 Identities for the trigonometric functions

The most important formulae to remember are

$$\sin^2 A + \cos^2 A = 1 \tag{1.6}$$

$$\cos(A+B) = \cos A \cos B - \sin A \sin B \tag{1.7}$$

$$\sin(A+B) = \sin A \cos B + \cos A \sin B. \tag{1.8}$$

If you have trouble remembering which of the last two is which, and which has the minus in it, try substituting some special values such as $A = 0$ or $B = \frac{1}{2}\pi$ and checking the result. For example, taking $A = 0$ in the last equation gives $\sin B = 0 + \sin B$, consistent, whereas if you had tried $\sin(A+B) = \sin A \cos B - \cos A \sin B$ you would get $\sin B = 0 - \sin B$, clearly wrong. From these and the earlier results Eq. 1.1 we get

$$\begin{aligned} \cos(A-B) &= \cos A \cos B + \sin A \sin B \\ \sin(A-B) &= \sin A \cos B - \cos A \sin B. \end{aligned}$$

and by adding or subtracting various pairs of the above equations, we get

$$\cos A \cos B = \tfrac{1}{2}(\cos(A+B) + \cos(A-B)) \tag{1.9}$$

$$\sin A \sin B = \tfrac{1}{2}(\cos(A-B) - \cos(A+B)) \tag{1.10}$$

$$\sin A \cos B = \tfrac{1}{2}(\sin(A+B) + \sin(A-B)), \tag{1.11}$$

which we will find very useful in doing integrations like $\int \cos(nx)\cos(mx)\,\mathrm{d}x$ which turn up later on.

The double angle cases

$$\begin{aligned}\sin 2x &= 2\sin x\cos x\\ \cos 2x = \cos^2 x - \sin^2 x = 2\cos^2 x - 1 &= 1 - 2\sin^2 x\\ \cos^2 x &= \tfrac{1}{2}(1+\cos 2x)\\ \sin^2 x &= \tfrac{1}{2}(1-\cos 2x)\end{aligned}$$

come up often; we get the "half angle" cases by just substituting in $y = 2x$, $x = y/2$ in the above.

Using the basic identities we can easily derive plenty more, such as

$$\begin{aligned}\sec^2 A &= 1 + \tan^2 A\\ \cos C + \cos D &= 2\cos\tfrac{1}{2}(C+D)\cos\tfrac{1}{2}(C-D).\end{aligned}$$

We should also note (see Thomas 3.4) that for any constant k,

$$\frac{\mathrm{d}(\sin(kx))}{\mathrm{d}x} = k\cos(kx), \quad \frac{\mathrm{d}(\cos(kx))}{\mathrm{d}x} = -k\sin(kx)\ .$$

Both $\sin(kx)$ and $\cos(kx)$ therefore obey[1]

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -k^2y\ .$$

and it can be proved that these give all solutions, i.e.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -k^2y \quad \Leftrightarrow \quad y = a\cos(kx) + b\sin(kx) \tag{1.12}$$

for some constants a and b. We could also write the right side as a combination of e^{ix} and e^{-ix}.

1.2 Ln, or $\log_e$, exp, and hyperbolic functions

(See Thomas section 7.2)
The natural logarithm $\ln x$ can be defined as

$$\ln x = \int_1^x \frac{\mathrm{d}t}{t}\ .$$

This implies $\ln 1 = 0$. Note that this is not a good definition if $x < 0$, but it is easy to show that for negative x, $\int^x \mathrm{d}u/u = \ln|x| + \text{constant}$. The number e (Euler's number) is then defined by $\ln e = 1$. (After π, this is the second most important constant in maths).

From the definition it is obvious that

$$\frac{\mathrm{d}\ln x}{\mathrm{d}x} = \frac{1}{x}\ .$$

One also finds:

$$\ln(ab) = \ln a + \ln b\ ,$$

Repeated application of this shows $\ln(a^n) = n\ln a$ for integer n, and it turns out this is true for any power p i.e.

$$\ln x^p = p\ln x,$$

[1] Those who have done applied maths. at A-level or later may recognize this as an equation for simple harmonic motion.

In particular, either putting $p=-1$ in the above, or $b=1/a$ in the previous equation, gives us

$$\ln(a^{-1}) = \ln(1/a) = -\ln(a)$$

and hence

$$\ln(a/b) = \ln a - \ln b\ .$$

Note that $\ln(a+b) \neq \ln a + \ln b$ (unless $a+b=ab$).

We can define exp (see Thomas 7.3) to be the inverse function to ln, so that $\exp(\ln x) = \ln(\exp x) = x$. Then $\exp 1 = e$ and $\exp r = e^r$. Note that $\exp(a+b) = e^a e^b$ NOT $e^a + e^b$. For any number a, $a = e^{\ln a}$ and hence $a^x = (e^{\ln a})^x = e^{x \ln a}$. In particular, this enables us to relate the usual logarithms (base 10) to natural logarithms since if $x = \log_{10} y$, $y = 10^x = e^{x \ln 10}$, so $\ln y = x \ln 10$ and $x = \ln y / \ln 10$. For a general a we can define $\log_a y = x$ to be such that $y = a^x$, so $\ln y = \log_e y$.

One can show that

$$\frac{\mathrm{d}\exp x}{\mathrm{d}x} = \exp x\ ;$$

to prove that, take $y = \exp x$, take ln of both sides so $\ln y = x$, then differentiate giving $(1/y)dy/dx = 1$ (by the chain rule), and rearrange.

We can now use e^x to define the hyperbolic functions (see Thomas 7.8)

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}), \quad \sinh x = \tfrac{1}{2}(e^x - e^{-x})\ .$$

These functions have identities and derivative properties that run closely parallel to those of sin and cos. If you know the trigonometric identities, the identities for hyperbolic functions can be recovered by substituting cosh for cos and i sinh for sin, where $i^2 = -1$.

From differentiating e^x we find

$$\frac{\mathrm{d}\sinh kx}{\mathrm{d}x} = k\cosh kx, \quad \frac{\mathrm{d}\cosh kx}{\mathrm{d}x} = k\sinh kx\ .$$

Thence

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = k^2 y \Leftrightarrow y = a\cosh(kx) + b\sinh(kx) \tag{1.13}$$

for some constants a and b (we can also write y as a combination of e^x and e^{-x}).

Comparing Eq. 1.13 with 1.12, we now see how to solve $d^2y/dx^2 = Cy$ for any constant C : if C is positive, we define $k = \sqrt{C}$ and get 1.13, while if C is negative we define $k = \sqrt{-C}$ and get 1.12; finally if $C = 0$ we easily integrate twice to get $y = ax + b$.

1.3 Double and triple integrals

(See Thomas 15.1 and 15.4)
First let us revise the idea of 2-D integration.

Example 1.1. Integrate the function $f(x,y) = x^2y^2$ over the triangular area $\mathscr{R}$: $0 \le x \le 1$, $0 \le y \le x$.

We can write this integral as

$$\int_{\mathscr{R}} f(x,y)\,\mathrm{d}A,$$

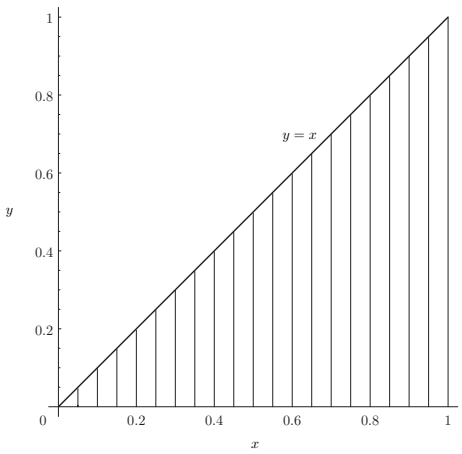

Figure 1.1: Integrating over the triangular region $\mathscr{R} : 0 \le x \le 1, 0 \le y \le x$.

where dA is an area element. But the area of a little rectangle of length δx in the x-direction and length δy in the y-direction is $\delta A = \delta x\,\delta y$; hence we can rewrite d$A$ as d$A = \mathrm{d}x\,\mathrm{d}y$. Thus the integral we want (cf. Fig. 1.1) is

$$\begin{aligned}
\int\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_{x=0}^{1}\left(\int_{y=0}^{x} x^2y^2\,\mathrm{d}y\right)\mathrm{d}x \\
&= \int_{x=0}^{1}\left(x^2\int_{y=0}^{x} y^2\,\mathrm{d}y\right)\mathrm{d}x \\
&= \int_{0}^{1} x^2\left(\frac{1}{3}x^3\right)\mathrm{d}x \\
&= \left[\frac{1}{18}x^6\right]_0^1 = \frac{1}{18}.
\end{aligned}$$

Here there are two key points to note: the limits on the (inner) y-integral depend on x, and in the second step we have moved the x^2 outside the y integral because it does not depend on y; so, the x^2 behaves like a "constant" inside the y-integral , but not for the x integral.

An area integral such as this is often called a double integral (because it can be rewritten as two 1-D integrations). Some authors use two integration signs, to remind you that it is an area integral: thus they would write $\int\int f(x,y)\,\mathrm{d}A$. In this course, **whenever it is obvious that an integral is over area, we shall generally just write** $\int f(x,y)\,\mathrm{d}A$.

Similarly some books write $\int\int\int f(x,y,z)\,\mathrm{d}V$ for a volume integral: **where no confusion will arise, we shall just write** $\int f(x,y,z)\,\mathrm{d}V$.

We shall need to put in all the integral signs when obtaining a value by doing the two or three integrations with respect to coordinates.

Exercise 1.1. Calculate $\int_{\mathscr{R}} f(x,y)\,\mathrm{d}A$ for

$$f(x,y) = 1 - 6x^2y \qquad \text{and} \qquad \mathscr{R} : 0 \le x \le 2, -1 \le y \le 1.$$

[Answer: 4] □

Note that in that exercise, the region of integration is a rectangle, so the limits of both the x- and y-integrations were constants so one could do the x- or the y-integration first – the answer will be the same. This holds for any rectangular region in 2-D , or for a cuboid when we come to 3-D integration.

In example 1.1, we had a triangle; now the upper limit of the "inside" integral depends on the "outer" variable. upper limit of the y-integral was x, so the y-integration had to be performed first with the limits as given. Otherwise the answer would have read

$$\int\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y \quad = \quad \int_{y=0}^{x}\left(\int_{x=0}^{1} x^2y^2\,\mathrm{d}x\right)\mathrm{d}y \;=\frac{x^3}{9}.$$

This depends on x, which is ridiculous as the answer is for a whole area, not some value of x. If we want to change the order of integration we need to take y going from 0 to 1, then x runs from y to 1 (check the sketch); now we have to put the dx integral on the inside, and we get

$$\int\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y \quad = \quad \int_{y=0}^{1}\left(\int_{x=y}^{1} x^2y^2\,\mathrm{d}x\right)\mathrm{d}y.$$

Check that this does give the same answer.

It's important that you understand how to get these limits: when doing a numerical evaluation of a multiple integral, there are several rules to remember :

i) Work out the limits on each variable from a sketch.

ii) The limits on each integral may depend on the variables x,y etc appearing as dx,dy **outside** that integral, but should not depend on those inside it. So the limits on the outermost integral sign should not depend on any of x,y,z; if dx is the outermost integral

iii) The limits on each integral apply to the "matching" variable, again working from inside to outside. So the last integral sign matches the first one of dx,dy,dz etc.

iv) Evaluate the resulting multiple integral from the "inside out", so you evaluate the innermost integration first. Putting in brackets can be helpful here, as in the example above.

It is a straightforward step from double integrals to volume integrals (triple integrals) of the form $\int_V f(x,y,z)\,\mathrm{d}V$. In Cartesian coordinates we have $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ (the volume of a 3-D rectangular box) and so

$$\int_V f(x,y,z)\;\mathrm{d}V = \int\int\int_V f(x,y,z)\;\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

Sometimes the geometry of the volume will make other choices of coordinate system preferable. In Thomas 15.3 and 15.6, which were studied in Calculus II, two-dimensional plane integrals in polar coordinates, and triple integrals in spherical and cylindrical polars are discussed: you will find it very useful to revise those sections. For a general change of coordinate system from Cartesians (x,y,z) to (u,v,w),

$$\int\int\int_V f\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int\int\int_V f\,J\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w$$

where J is the Jacobian determinant of (x,y,z) with respect to (u,v,w). This determinant J is the volume ratio of the two coordinate systems: if we take an infinitesimal cuboid in (u,v,w) space of volume $du\,dv\,dw$, this will map to a parallelepiped in x,y,z space, and J is the ratio of those volumes (if you need to revise this in more detail, see Thomas 15.7).

Exercise 1.2. Evaluate $\int\int\int e^x\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ over the volume V of the tetrahedron bounded by the four planes $x=0$, $y=0$, $z=0$ and $x+y+z=a$ ($a>0$). [Answer: $e^a-\frac{1}{2}a^2-a-1$.] □

1.4 Curves and surfaces

We shall use various geometrical shapes in examples, so we need the equations for them. The main ones are so-called 'conic sections' in two dimensions, and related three-dimensional surfaces. Other courses also discuss more complicated shapes (see e.g. Thomas 10.6 and 10.7).

First we discuss curves in two dimensions. There are three main ways to specify a curve:

One way is to give an equation $y = f(x)$; a second way is to give an equation $g(x,y) = 0$: the curve is then the set of points (x, y) obeying the equation. Given the first form, we can get the second by defining $g(x,y) = y - f(x)$, but not necessarily the converse

A third way is the **parametric form** in terms of two functions of some variable t: $x = a(t),\ y = b(t)$ (see Thomas 3.5). Sometimes we can take $t = x$ itself. The parametrized form carries extra information, about which direction and how fast we go along the curve as t changes.[2] We will see a lot more examples of the parametrised form in Chapter 2

Using the second way, some standard curves are:

$$x^2 + y^2 = a^2 \qquad \text{circle, centre } (0,0), \text{ radius } a \tag{1.14}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad \text{ellipse, centre } (0,0), \text{ semi-major axes } a \text{ and } b \tag{1.15}$$

$$y = ax^2 + b \qquad \text{parabola, symmetric about } x = 0 \tag{1.16}$$

$$cy^2 - kx^2 = a^2, (ck > 0) \qquad \text{hyperbola, symmetric about } x = 0 \text{ and } y = 0 \tag{1.17}$$

(See Thomas 1.2, 1.5.) The special case of a hyperbola with $a = 0$ is just a pair of straight lines. These curves involving only constants and powers up to x^2 and y^2 are known as the **conic sections**.

To recognize these, first look for the coefficients of the x^2 and y^2:
if one is 0, but the corresponding variable appears linearly, it's a parabola;
if they have the same sign it's an ellipse (or as a special case a circle), and
if they have opposite sign it's a hyperbola
(assuming the remaining constants allow there to be some points: $x^2 + y^2 = -5$ has no real points).

What if the equation is quadratic but not one of these standard forms? Given

$$x^2 + 6x + y^2 + 8x = 0$$

we can carry out a process called 'completing the square' to write it as

$$(x+3)^2 + (y+4)^2 = 25$$

which we now recognize as a circle radius 5, centre $(-3, -4)$: this circle passes through the origin. Similar methods can be used to recognize the other standard curves if they are given relative to origins different from the ones used in the most standard forms below (cf. Thomas 1.5).

We can also recognize the case where the axes have been transformed, in a similar way. For example, $xy = b^2 \Leftrightarrow (y+x)^2 - (y-x)^2 = 4b^2$, so it's a hyperbola where the symmetry axes are at 45° to those used in (1.17) with $c = k = 1$ and with $4b^2 = a^2$. In general we have to complete the square on the terms quadratic in x and y: for example the rearrangement

$$x^2 + 4xy + 3y^2 = (x+2y)^2 - y^2$$

shows the curve $x^2 + 4xy + 3y^2 = 6$ is not an ellipse, as you might think from the fact the coefficients of x^2 and y^2 are both positive, but a hyperbola.

[2] The latter approach is used heavily in Geometry II.

Parametrized curves are also useful, especially when calculating **line integrals** along curves later on. Here are some standard parametrizations for the circle, ellipse and hyperbola:

$$x^2+y^2=a^2 \qquad (x,y)=(a\cos\theta, a\sin\theta) \tag{1.18}$$

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \qquad (x,y)=(a\cos\theta, b\sin\theta) \tag{1.19}$$

$$y^2-x^2=a^2 \qquad (x,y)=(a\sinh\theta, a\cosh\theta). \tag{1.20}$$

These work because of the identity (1.6) and its hyperbolic counterpart $\cosh^2 x-\sinh^2 x=1$. (See Thomas 10.4 for futher or alternative parametrizations.) We shall use these, especially the first two, later.

1.5 Surfaces in 3-D

[Here we meet material you may not have seen before.]

For surfaces in 3 dimensions, there are similarly three main ways to give the equations. One is to give one coordinate in terms of the other two, e.g. $z=h(x,y)$. Another is to use a single equation $V(x,y,z)=0$. The third is by a parametrization in terms of two variables e.g. $(x(u,v),y(u,v),z(u,v))$ (see Thomas 16.6, and more details in Chapter 2).

We shall again focus on surfaces described by quadratics in x, y and z at worst. To work out what the surface is like, one good way is to consider letting one coordinate be constant, for example $z=d$, which means we are considering a "slice" through the surface $V=0$ at the plane $z=d$. The intersection of a curved surface and a plane is generally a 1-D curve, which we should be able to identify from the previous section. Then we just stack those curves for varying d.

One simple case is

$$x^2+y^2=a^2$$

The equation is the same as for a circle, but as we are now in 3 dimensions, it implies z can take any value. In each plane $z=d$ we have a circle. Hence, this is an infinite circular cylinder along the z-axis. Very often some bounding values of z are given, e.g. $0\le z\le 2$. Then we have a finite cylinder, the shape of a drinks can.

Example 1.2. What is the surface $\frac{y^2}{a^2}+\frac{z^2}{b^2}=1$?

It is an infinite elliptical cylinder along the x axis.

We can also have parabolic and hyperbolic "cylinders", using (1.16) and (1.17).

Another simple three-dimensional surface is that of a **sphere** of radius a centred at the origin:

$$x^2+y^2+z^2=a^2\ . \tag{1.21}$$

Example 1.3. What is the surface $x^2+y^2+z^2=a^2$, $x\ge 0$?

The hemisphere to the right of the plane $x=0$.

We can put together cases where we get one of the standard types of curve listed earlier in planes $z=d$ and different ones in planes $x=k$ or $y=m$ say.

For example, we can generalize the ellipse (1.15) to

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1 \tag{1.22}$$

(see Thomas, 12.6). In this case each of the three types of cut-plane gives an ellipse as the curve. The surface is an *ellipsoid*. This shape, and the ones that follow, are shown in diagrams 12.48-12.52 in Thomas. (also, the Wikipedia article on "Quadrics" has some pretty graphics).

If instead we had taken

$$\frac{z}{c}=\frac{x^2}{a^2}+\frac{y^2}{b^2}\ , \tag{1.23}$$

we then have an ellipse in each plane $z=d$ but a parabola in each plane $x=k$ or $y=m$. This is an *elliptic paraboloid*. Changing the plus to a minus in this equation gives a hyperbolic paraboloid.

Similarly we can obtain an *(elliptic) hyperboloid* as

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=+1\ . \tag{1.24}$$

Here we have ellipses in planes $z=d$, and hyperbolae in the planes $x=0$ and $y=0$. Moving the z^2 term over to the RHS, we see the RHS is positive for any z, so there is an ellipse for any fixed value of z and the surface has just one piece (we say 'one sheet').

However, if instead we had a -1 on the right, i.e.

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1\ . \tag{1.25}$$

we can rearrange into

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{z^2}{c^2}-1\qquad . \tag{1.26}$$

It's now clear that if $z^2/c^2>1$, i.e. $z<-c$ or $z>c$), we again get an ellipse in the xy plane; but if $-c<z<c$ the RHS is negative and there are no solutions for x,y. This is a hyperboloid of two sheets.

The elliptic paraboloid and hyperboloid have circular special cases where $a=b$. Note also that we can swap x, y and z around in these forms so we have different choices of axes for the same shapes.

There is also the special case of the hyperboloid equation where the constant on the RHS is zero, i.e.

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0 \tag{1.27}$$

This is a cone through the origin. Taking a plane through the origin such as $x=0$, we get two straight lines, while taking planes perpendicular to the axes but not through the origin gives ellipses or parabolae. In fact all the quadratic curves (ellipses, circles, parabolae and hyperbolae) can be obtained by intersecting the circular cone $z^2=x^2+y^2$ with planes (not necessarily perpendicular to the axes): this is why they are called conic sections (see Thomas chapter 10).

If we are given a quadratic surface in a different form, we can first rearrange it into one of the forms above: rearrange so that all the x,y,z terms are on the left, and the constant on the right; if the constant is not zero, divide by it to get a $+1$ on the right; then look at the x^2, y^2 and z^2 parts: if two or three of these have negative coefficients, just multiply by -1 to make at least two of the coefficients of x^2, y^2, z^2 positive, Then,

If all 3 are positive, it's an ellipsoid (or a sphere).

If two are positive and one negative it's a hyperboloid, and we need to check the constant term

to see if it's one sheet or two

If one of the three x^2,y^2 or z^2 terms is zero, but there is a linear term in the corresponding variable, it's a paraboloid: the relative sign of the other two show if it is elliptic or hyperbolic.

If one variable is missing completely, it's a "cylinder" given by the matching 2-D curve.

As in the case of curves, we can work out what the shape is if the equations are not in standard form but have shifted origins or rotated axes, by completing the square. For this course, we'll keep it simple though, so we will only be looking at surfaces which are aligned with the coordinate axes.

Now for some parametrized versions (see Thomas 16.6)

	Implicit form	Parametric form		
Cylinder	$x^2+y^2=a^2$	$(x,y,z)=(a\cos\theta, a\sin\theta, z)$	Parameters θ, z	(1.28)
Sphere	$x^2+y^2+z^2=a^2$	$(a\sin\theta\cos\phi, a\sin\theta\sin\phi, a\cos\theta)$	Parameters θ, ϕ	(1.29)
Ellipsoid	$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=+1$	$(a\sin\theta\cos\phi, b\sin\theta\sin\phi, c\cos\theta)$	Parametersθ, ϕ	(1.30)
Hyperboloid	$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$	$(a\cos u, b\sin u\cosh v, c\sin u\sinh v)$	Parameters u, v	(1.31)

1.6 Vectors

(Note: this is in chapter 12 in Thomas but in Geometry I you used Hirst)
Vectors can be introduced as displacements in space, called **position vectors**. To describe a position vector, we need to specify its direction and its length or magnitude (to say how far we go in the given direction). This is a geometric definition. A vector is different from a **scalar**, a quantity which has only a magnitude but no direction.

One can draw a vector as an arrow of the appropriate length and direction. Vectors are usually notated in print by boldface type, e.g. **a**, and in handwriting by under- or over-lining such as $\underline{a}$, $\vec{a}$, or $\underset{\sim}{a}$.

Warning: When writing, it is tempting to miss off the under/overlines to save time. This is a bad idea, because if you confuse what's a scalar and what's a vector in your working, you immediately get nonsense.

To define a vector algebraically, i.e. in a formula, we can use the Cartesian coordinates of the point to which it displaces the origin, e.g.

$$\mathbf{r}=(x,y,z). \tag{1.32}$$

Note: As you saw in Geometry I, we can write vectors either as row or column vectors. The column vector form is useful if you are multiplying by matrices (like rotation matrices), but in this course we shall mainly use the row vector form which is more compact.)

Here x, y and z are called the *components* of **r**. We may refer to (x,y,z) as the point **r**. From now on we shall use the notation **r** only for this vector.

The length of a vector **v** is denoted by $|\mathbf{v}|$ or sometimes just v; this is a scalar. The vector **r** has length $r=\sqrt{x^2+y^2+z^2}$, by Pythagoras' theorem in 3 dimensions.

To add vectors **a** and **b** we simply take the displacement obtained by displacing first by **a** and then by **b** (the result can be defined as the diagonal of the parallelogram with sides **a** and **b**). In components this says that $\mathbf{v}=(v_1,v_2,v_3)$ and $\mathbf{w}=(w_1,w_2,w_3)$ have the sum

$$\mathbf{v}+\mathbf{w}=(v_1+w_1, v_2+w_2, v_3+w_3).$$

Subtraction can then be defined similarly. The zero vector $\mathbf{0}$ is the one with zero magnitude (and no well-defined direction!).

It is now easy to show this obeys the usual rules of addition (and subtraction).[3]

The displacement from a point $\mathbf{r}_1$ to a point $\mathbf{r}_2$ is $\mathbf{r}_2 - \mathbf{r}_1$.

We can multiply a vector by a scalar (a number) λ, simply by multiplying its magnitude, preserving the direction. In components, if $\mathbf{v} = (v_1, v_2, v_3)$ then we have

$$\lambda \mathbf{v} = (\lambda v_1, \lambda v_2, \lambda v_3).$$

This operation also obeys very simple and obvious rules.[4] This multiplication gives us a way to define the unit vector (the vector of length 1) in the same direction as $\mathbf{v}$, denoted by $\hat{\mathbf{v}}$, by $\hat{\mathbf{v}} \equiv \mathbf{v}/|\mathbf{v}|$ (strictly, we should write the number first so we would have to write $(1/|\mathbf{v}|)\mathbf{v}$, but in practice it's obvious what we mean).

These rules give us another common way of writing a vector. We note that we can arrive at the same total displacement by first moving along the x-axis, then parallel to the $y-$axis then parallel to the $z-$axis; and we can express this by defining the unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ along the directions of the three axes by

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

This way of writing (1.32) has the advantage of making it clearer how the components change if we change our choice of axes: if we rotate our axes to a different system x', y', z', we will get 3 new unit vectors e.g. **i'**, **j'** and **k'**, and converting vectors between systems looks like a matrix multiplication - more on this later.

Note that all of these statements about position vectors in 3 dimensions can very simply be applied in 2 dimensions also, with obvious minor changes.

Although we have motivated vectors by introducing them as displacements, they can represent, or be interpreted as, many other things: for example, a force, a velocity, inputs and outputs in an economic model, and so on.

A parametric equation of the type

$$\mathbf{r} = \mathbf{p} + t\mathbf{q}, \qquad -\infty < t < \infty \tag{1.33}$$

defines a line through point $\mathbf{p}$ parallel to direction $\mathbf{q}$. For example $\mathbf{r} = t\mathbf{k},\ -\infty < t < \infty$ is the z axis.

Using this, we can get the straight line going through two given points $\mathbf{r}_1$ and $\mathbf{r}_2$: the vector from $\mathbf{r}_1$ to $\mathbf{r}_2$ is $\mathbf{r}_2 - \mathbf{r}_1$, so the (infinite) line through them is

$$\mathbf{r} = \mathbf{r}_1 + t(\mathbf{r}_2 - \mathbf{r}_1), \quad infty < t < \infty \tag{1.34}$$

If instead we take a range $0 <\leq t \leq 1$ in the above, this gives us the finite line segment with end-points at the two given points. This will be very useful later on, *memorise it*.

[3] This means that for any vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$,

$$\mathbf{a}+\mathbf{b} = \mathbf{b}+\mathbf{a}, \quad (\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c}), \quad \exists \mathbf{0} \text{ such that } \mathbf{a}+\mathbf{0} = \mathbf{a},$$

and given $\mathbf{a}$, $\exists(-\mathbf{a})$ such that $\mathbf{a}+(-\mathbf{a}) = \mathbf{0}$. These rules are purely abstract and make no reference to displacements or three dimensions, and are part of the general definition of a vector space which is given in Linear Algebra I. Those who have encountered groups will recognise that they ensure that the space of vectors is an additive group under vector addition.

[4] More precisely, for any vectors $\mathbf{a}$ and $\mathbf{b}$, and numbers λ and μ, we have

$$\lambda(\mathbf{a}+\mathbf{b}) = \lambda\mathbf{a}+\lambda\mathbf{b}, \quad (\lambda+\mu)\mathbf{a} = \lambda\mathbf{a}+\mu\mathbf{a}, \quad (\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a})$$

and $1\mathbf{a} = \mathbf{a}$. For a general vector space, as defined in Linear Algebra I, the scalars are elements of a general field but here we shall only use the real numbers $\mathbb{R}$. However, these rules do apply when λ and μ are elements of a general field, for instance the complex numbers $\mathbb{C}$.

Example 1.4. Medians of a triangle

Vectors can often be used to derive geometrical results very concisely, as this example shows.

Let $\mathbf{a},\mathbf{b},\mathbf{c}$ be the corners of a triangle. The midpoint of the side connecting $\mathbf{b}$ and $\mathbf{c}$ will be $\frac{1}{2}(\mathbf{b}+\mathbf{c})$. A line through this midpoint and $\mathbf{a}$ is

$$\mathbf{r} = \mathbf{a} + t(\tfrac{1}{2}\mathbf{b} + \tfrac{1}{2}\mathbf{c} - \mathbf{a}), \qquad -\infty < t < \infty$$

which is called the median through $\mathbf{a}$. Putting $t = \frac{2}{3}$ (note: here this choice is a rabbit out of the hat, but we can find it by writing down a second median and solving for the intersection point) we get the point $\frac{1}{3}(\mathbf{a}+\mathbf{b}+\mathbf{c})$. Since this point is symmetric in $\mathbf{a},\mathbf{b},\mathbf{c}$, the medians through $\mathbf{b}$ and $\mathbf{c}$ will also pass through it. Hence the three medians of a triangle intersect at a single point.

If we write out the components of (1.33), with notation $\mathbf{r} = (x, y, z)$, $\mathbf{p} = (p_1, p_2, p_3)$, $\mathbf{q} = (q_1, q_2, q_3)$ we find

$$x = p_1 + tq_1 \quad y = p_2 + tq_2, \quad z = p_3 + tq_3,$$

from which we can eliminate t to get

$$\frac{x - p_1}{q_1} = \frac{y - p_2}{q_2} = \frac{z - p_3}{q_3},$$

giving the two independent linear equations (e.g. for y and z in terms of x) needed for a line in three-dimensional space.

We can now write functions of 3-dimensional position $f(x,y,z)$ more compactly as functions $f(\mathbf{r})$. Equations of the form $f(\mathbf{r}) =$ constant define surfaces, the constant surfaces of f. A simple example is $r^2 = 1$, which is a sphere of unit radius centred at the origin. (Recall our notation allows $r \equiv |\mathbf{r}|$.)

Example 1.5. A sphere

The geometrical interpretation of

$$|\mathbf{r}-\mathbf{k}|=1$$

as a sphere of unit radius centred at $(0,0,1)$ is obvious. Equivalent expressions are $x^2+y^2+(z-1)^2=1$ and $x^2+y^2+z^2-2z=0$.

Warning: One of the commonest errors made by students is to confuse vectors and scalars, in particular to start adding together the components of a vector. The vector $(3,1,2)$ is not the same as the scalar 6. This may seem obvious now, but the mistake is more easily made when using basis vectors like **i**, **j** and **k**; then it somehow seems to be easier to make the mistake $3\mathbf{i}+\mathbf{j}+2\mathbf{k}=6$.

1.7 Scalar and vector products

We have defined vector addition and subtraction, but not multiplication of vectors. This is more complicated because to obtain another vector we need to define both a magnitude and a direction (and in general, vector division cannot be defined at all; we can divide a vector by a scalar λ just by multiplying by $1/\lambda$, but we cannot divide anything by a vector).

We first define the dot product, or scalar product[5], whose result is not a vector but a scalar. For vectors **v** and **w**, this is defined by

$$\mathbf{v}.\mathbf{w}\equiv|\mathbf{v}||\mathbf{w}|\cos\theta, \tag{1.35}$$

where θ is the angle between **v** and **w**. An alternative definition in terms of the components (v_1,v_2,v_3) and (w_1,w_2,w_3) of **v** and **w** is

$$\mathbf{v}.\mathbf{w}\equiv v_1w_1+v_2w_2+v_3w_3=\sum_{i=1}^{3}v_iw_i.$$

One can prove that the two definitions are the same by applying Pythagoras' theorem to a triangle constructed as follows. Take sides **v**, **w** and $\mathbf{v}+\mathbf{w}$. Draw the perpendicular from $\mathbf{v}+\mathbf{w}$ to the line in direction **v**. It has height $|\mathbf{w}|\sin\theta$ and meets the direction **v** at a distance $|\mathbf{v}|+|\mathbf{w}|\cos\theta$. Now write out Pythagoras with the lengths in terms of $|\mathbf{v}|$, $|\mathbf{w}|$ and θ and again in terms of components and compare the results. The details are left as an exercise (if you have trouble, look in the online notes for MAS114 Geometry I or in A.E. Hirst, Vectors in 2 or 3 dimensions, Arnold 1995, chapter 3).

We note in particular that two non-zero vectors **v** and **w** are perpendicular (θ is a right angle) if and only if $\mathbf{v}.\mathbf{w}=0$.

Example 1.6. (This example was used in Geometry I.)
Find $\cos\theta$ where θ is the angle between $\mathbf{v}=(1,3,-1)$ and $\mathbf{w}=(2,2,1)$.

$$\begin{aligned}\mathbf{v}.\mathbf{w} &= 1.2+3.2+(-1).1=7=|\mathbf{v}||\mathbf{w}|\cos\theta,\\ |\mathbf{v}|^2 &= 1^2+3^2+(-1)^2=11,\\ |\mathbf{w}|^2 &= 2^2+2^2+1^2=9,\ \text{so}\\ \cos\theta &= \frac{7}{\sqrt{11}\sqrt{9}}=\frac{7}{3\sqrt{11}}\end{aligned}$$

[5] In a more abstract setting (such as in Linear Algebra I) this may also be called the inner product.

From either form of the definition we can easily derive various algebraic rules.[6]

A geometrical application of the dot product is in giving the equation of a plane. The plane through a fixed point $\mathbf{p}$ perpendicular to a fixed vector $\mathbf{v}$ is given by the set of all points $\mathbf{r}$ which have $\mathbf{r}-\mathbf{p}$ perpendicular to $\mathbf{v}$, as is easily seen from a sketch. Since two perpendicular vectors have a dot product of zero, this gives

$$(\mathbf{r}-\mathbf{p}).\mathbf{v}=0 \tag{1.36}$$

This easily rearranges to $\mathbf{r}.\mathbf{v}=\mathbf{p}.\mathbf{v}$ and the right-hand side is just a constant for given $\mathbf{p},\mathbf{v}$.

In components, if $\mathbf{v}=(a,b,c)$ and $\mathbf{p}.\mathbf{v}=d$ the equation for a plane reads $ax+by+cz=d$. In practice people often choose a unit vector $\mathbf{n}$ when specifying a plane in this form, so that $\mathbf{p}.\mathbf{n}$ becomes the perpendicular distance of the plane from the origin. Then, the distance of any other point $\mathbf{r}_1$ from that plane is given by $(\mathbf{r}_1-\mathbf{p}).\mathbf{n}=\mathbf{r}_1.\mathbf{n}-\mathbf{p}.\mathbf{n}=\mathbf{r}_1.\mathbf{n}-d$ (the sign here tells one which side of the plane $\mathbf{r}_1$ is on).

The vector product: To define a product of two vectors which is a third vector, we need to define a direction from two vectors $\mathbf{u}$ and $\mathbf{v}$. The only way to do this which treats the two vectors equally is to take the perpendicular to the plane in which $\mathbf{u}$ and $\mathbf{v}$ lie. However, this does not fully define a direction, because we need to know which way to go along the perpendicular. For that the convention is to use the so-called right-hand rule: hold the fingers of your right hand so they curl round from $\mathbf{u}$ to $\mathbf{v}$ and then take the direction your thumb points (see Thomas figures 12.27 and 12.28). If you do DIY, you may find it helpful to remember that this is the direction a normal screw travels if you turn your screwdriver clockwise. Note that this definition only works in **three dimensions**: there is no well-defined vector product in n dimensions for $n>3$.

The magnitude of $\mathbf{v}\times\mathbf{w}$ is defined to be $|\mathbf{v}||\mathbf{w}|\sin\theta$ (θ as before). Geometrically this is the area of a parallelogram with sides $\mathbf{v}$ and $\mathbf{w}$. Note that for perpendicular vectors this rule implies that the magnitude is $|\mathbf{v}||\mathbf{w}|$. These rules have the consequences that for any vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ and any scalar λ,

$$\begin{aligned}
\mathbf{v}\times\mathbf{w} &= -\mathbf{w}\times\mathbf{v},\\
(\lambda\mathbf{v})\times\mathbf{w} &= \lambda(\mathbf{v}\times\mathbf{w})=\mathbf{v}\times(\lambda\mathbf{w}),\\
\mathbf{u}\times(\mathbf{v}+\mathbf{w}) &= (\mathbf{u}\times\mathbf{v})+(\mathbf{u}\times\mathbf{w}),\\
(\mathbf{u}+\mathbf{v})\times\mathbf{w} &= \mathbf{u}\times\mathbf{w}+\mathbf{v}\times\mathbf{w},
\end{aligned}$$

and $\mathbf{v}\times\mathbf{w}=\mathbf{0}$ for non-zero $\mathbf{v},\mathbf{w}$ if and only if $\mathbf{v}$ and $\mathbf{w}$ are parallel or anti-parallel (in particular, $\mathbf{v}\times\mathbf{v}=\mathbf{0}$ for any $\mathbf{v}$).

Note: it is particularly important to note the sign-change property that $\mathbf{v}\times\mathbf{w}=-\mathbf{w}\times\mathbf{v}$. This looks "silly", but is a consequence of the "handedness" of three-dimensional space, and which way round we choose to label our three coordinate axes.

From the notation used, the vector product is often called the cross product.[7]

[6] The main ones are that $\mathbf{v}.\mathbf{w}=\mathbf{w}.\mathbf{v}$ and that for any vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ and any scalar λ,

$$\begin{aligned}
\mathbf{v}.(\lambda\mathbf{w}) &= \lambda(\mathbf{v}.\mathbf{w})=(\lambda\mathbf{v}).\mathbf{w},\\
\mathbf{u}.(\mathbf{v}+\mathbf{w}) &= (\mathbf{u}.\mathbf{v})+(\mathbf{u}.\mathbf{w}),\\
(\mathbf{u}+\mathbf{v}).\mathbf{w} &= (\mathbf{u}.\mathbf{w})+(\mathbf{v}.\mathbf{w}),\\
\mathbf{v}.\mathbf{v} &= |\mathbf{v}|^2\geq 0,\\
\mathbf{v}.\mathbf{v} &= 0\Leftrightarrow\mathbf{v}=\mathbf{0}.
\end{aligned}$$

[7] You will also find that in some texts it is denoted $\mathbf{v}\wedge\mathbf{w}$, but I strongly advise against using this notation as it leads to confusion in more general settings where $\mathbf{v}\wedge\mathbf{w}$ is not a vector. The reason for this misuse is that $\mathbf{v}\wedge\mathbf{w}$ is what's called a two-form, and there is an operation called the Hodge dual, denoted by $*$, such that in three dimensions $*(\mathbf{v}\wedge\mathbf{w})=\mathbf{v}\times\mathbf{w}$.

To get the expressions for the cross product in terms of components, we can start by noting that the unit vectors **i**, **j** and **k** are perpendicular to one another (so the vector product of any two distinct ones among them has magnitude 1). This means that $\mathbf{i}\times\mathbf{j}$ must have length 1 and be perpendicular to both of them, so it is either $+\mathbf{k}$ or $-\mathbf{k}$. Since the usual x, y and z axes, in that order, are a right-handed set, it will turn out that

$$\mathbf{i}\times\mathbf{j}=+\mathbf{k},\quad \mathbf{j}\times\mathbf{k}=+\mathbf{i},\quad \mathbf{k}\times\mathbf{i}=+\mathbf{j},$$

and therefore

$$\mathbf{j}\times\mathbf{i}=-\mathbf{k},\quad \mathbf{k}\times\mathbf{j}=-\mathbf{i},\quad \mathbf{i}\times\mathbf{k}=-\mathbf{j}.$$

(To remember these , think of the sequence **ijkijk**...; if the two vectors in the cross-product are in the same order as in that sequence, the RHS has a $+$ sign, while if they are in reverse order there is a $-$ sign.)

Also $\mathbf{i}\times\mathbf{i}=\mathbf{j}\times\mathbf{j}=\mathbf{k}\times\mathbf{k}=\mathbf{0}$. Using these we easily obtain

$$(v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k})\times(w_1\mathbf{i}+w_2\mathbf{j}+w_3\mathbf{k})=(v_2w_3-v_3w_2)\mathbf{i}+(v_3w_1-v_1w_3)\mathbf{j}+(v_1w_2-v_2w_1)\mathbf{k}. \tag{1.37}$$

This can also be written as the formal determinant

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

One geometrical use of the cross product is in forming the volume of a parallellepiped with sides **u**,**v** and **w**. Thinking of (say) **u** and **v** as the base, and θ as the angle between $\mathbf{u}\times\mathbf{v}$ and **w**, so that the height is $|\mathbf{w}|\cos\theta$, we see that

$$\text{Volume of parallellepiped}=(\mathbf{u}\times\mathbf{v}).\mathbf{w} \tag{1.38}$$

(positive if **u**,**v** and **w** are a right-handed set). This quantity is called the *scalar triple product* and it is easy to show that

$$\mathbf{u}.(\mathbf{v}\times\mathbf{w})=\mathbf{v}.(\mathbf{w}\times\mathbf{u})=\mathbf{w}.(\mathbf{u}\times\mathbf{v})$$

(but this is $-\mathbf{v}.(\mathbf{u}\times\mathbf{w})$ etc, remember). We can also show that swapping the dot and cross gives the same result, i.e. $(\mathbf{u}\times\mathbf{v}).\mathbf{w}=\mathbf{w}.(\mathbf{u}\times\mathbf{v})=\mathbf{u}.(\mathbf{v}\times\mathbf{w})$ from above , but note that the brackets also move, i.e. the cross product must be done first (inside the brackets) otherwise the result is nonsense. (Some textbooks may omit the brackets, but this is potentially confusing). Clearly swapping the two vectors inside the bracket changes the sign, and we can show that this is also true for swapping any two of the three vectors.

Exercise 1.3. Prove from the definitions that, for all **a**, **b** and **c**,

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c})=(\mathbf{a}.\mathbf{c})\mathbf{b}-(\mathbf{a}.\mathbf{b})\mathbf{c}.$$

This quantity is called the **vector triple product**; note that the position of the brackets matters here.

□

1.8 Gradients and directional derivatives

(See Thomas 14.5 [and 16.2])
In Calculus II you met functions of more than one variable. Those that were discussed there were *scalar functions*, i.e. functions whose value at a particular point is a number. Such a scalar quantity (magnitude but no direction) that depends on position in space is called a **scalar field**. An example would be the temperature in a room – it has magnitude but not direction (so it is a scalar), and it is (in general) a function of position.

Suppose that $V(x,y,z)$ or $V(\mathbf{r})$ is a scalar field defined in some region. Then we can define a vector, the **gradient** of V, at each point, which we denote ∇V, as follows:

$$\nabla V = \frac{\partial V}{\partial x}\mathbf{i} + \frac{\partial V}{\partial y}\mathbf{j} + \frac{\partial V}{\partial z}\mathbf{k}.$$

So the x-, y- and z-components of the new vector are $\partial V/\partial x$, $\partial V/\partial y$ and $\partial V/\partial z$. See Thomas 14.5 if you need to revise this in more detail. Sometimes instead of ∇V we write grad V: the two notations are interchangeable.

Example 1.7. If $V(x,y,z) = x^2 \sin z$, calculate ∇V.

In this example, $\partial V/\partial x = 2x\sin z$, $\partial V/\partial y = 0$ and $\partial V/\partial z = x^2\cos z$. Hence

$$\nabla V = 2x\sin z\,\mathbf{i} + x^2\cos z\,\mathbf{k}.$$

Exercise 1.4. Evaluate the gradient ∇f of the following scalar fields.

(a) $f = x+y+z$,

(b) $f = yx^2 + y^3 - y + 2x^2z$,

(c) $f = \mathbf{a}.\mathbf{r}$, where $\mathbf{a}$ is a constant vector.

□

Now ∇V tells us how V changes if we move from one point to a nearby point. Suppose we start at a point $\mathbf{r} = (x,y,z)$, and then move a small distance $d\mathbf{r} = (dx,dy,dz)$ to the new point $\mathbf{r} + \mathrm{d}\mathbf{r} = (x+\mathrm{d}x, y+\mathrm{d}y, z+\mathrm{d}z)$: we will get a small change in V, given by

$$\begin{aligned} \mathrm{d}V &\equiv V(x+\mathrm{d}x, y+\mathrm{d}y, z+\mathrm{d}z) - V(x,y,z) \\ &= \frac{\partial V}{\partial x}\mathrm{d}x + \frac{\partial V}{\partial y}\mathrm{d}y + \frac{\partial V}{\partial z}\mathrm{d}z \end{aligned}$$

Here the second line uses the Taylor series in more than one variable, and discards terms in second and higher derivatives since $d\mathbf{r}$ is small. But $(\mathrm{d}x, \mathrm{d}y, \mathrm{d}z) = \mathrm{d}\mathbf{r}$, and so the right-hand side is just $\nabla V.\mathrm{d}\mathbf{r}$. Hence for a small change $d\mathbf{r}$, the change in V is

$$\mathrm{d}V = \nabla V.\mathrm{d}\mathbf{r} \tag{1.39}$$

Note: to use this, you **must** evaluate ∇V at the point concerned.

In our original definition of grad, it was implicitly assumed that we were working in terms of some specified Cartesian coordinate system (x,y,z). Equation (1.39) is important, because we can use it as a more

fundamental definition of grad, which will enable us to write down ∇V in more general coordinate systems. We shall return to this point later, in Chapter 5

Next, consider a surface

$$V(\mathbf{r}) = \text{constant},$$

and suppose that the point $\mathbf{r_1}$ is on that surface. If d$\mathbf{r}$ is a displacement *on* this surface, $V(\mathbf{r}_1 + \mathrm{d}\mathbf{r}) = V(\mathbf{r}_1)$. Thus $\mathrm{d}V = \nabla V|_{\mathbf{r}_1}.\mathrm{d}\mathbf{r} = 0$. Since this applies for every small displacement d$\mathbf{r}$ in the surface, $\nabla V|_{\mathbf{r}_1}$ must be perpendicular to the surface at $\mathbf{r}_1$. This gives us a way of finding a normal to a surface when the surface is specified by a single equation: the unit normal $\mathbf{n}$ will be $\nabla V/|\nabla V|$ evaluated at the point concerned.

Suppose we now want a normal line to $V =$ constant, or a tangent plane (see Thomas 14.6). As we know from (1.33) and (1.36), a line through point $\mathbf{a}$ in direction $\mathbf{q}$ can be written in parametric form as

$$\mathbf{r} = \mathbf{a} + t\mathbf{q}, \qquad -\infty < t < \infty,$$

while the plane through $\mathbf{a}$ perpendicular to $\mathbf{q}$ is

$$(\mathbf{r} - \mathbf{a}).\mathbf{q} = 0.$$

so all we have to do is insert values of $\mathbf{a}$ and $\mathbf{q}$ in these formulae. For the tangent plane and normal line to a surface at a given point $\mathbf{p}$, this gives

$$(\mathbf{r} - \mathbf{p}).\nabla V|_{\mathbf{p}} = 0 \qquad \text{and} \qquad \mathbf{r} = \mathbf{p} + t\nabla V|_{\mathbf{p}}.$$

It is sometimes convenient to eliminate the parameter t for the normal line, which we can do by taking the cross product with $\nabla V|_{\mathbf{p}}$:

$$(\mathbf{r} - \mathbf{p}) \times \nabla V|_{\mathbf{p}} = \mathbf{0}.$$

Note that the forms $(\mathbf{r} - \mathbf{p}).\mathbf{n} = 0$ and $(\mathbf{r} - \mathbf{p}) \times \mathbf{n} = \mathbf{0}$, using the unit normal $\mathbf{n}$, would give the same plane or line (though in $\mathbf{r} = \mathbf{a} + t\mathbf{n}$ such a change alters the values of t for given points) so we need not calculate $|\nabla V|$ to get the tangent plane or normal line.
Repeated Note: to use this, you **must** evaluate ∇V at the point concerned.

Exercise 1.5. Find equations for the (i) tangent plane and (ii) normal line at the point P_0 on each of the surfaces:

(a) $x^2 + 3yz + 4xy = 27, \qquad P_0 = (3, 1, 2).$

(b) $y^2z + x^2y = 7, \qquad P_0 = (2, 1, 3).$

[Answers: (a) $10x + 18y + 3z = 54$, $\mathbf{r} = (3 + 10t)\mathbf{i} + (1 + 18t)\mathbf{j} + (2 + 3t)\mathbf{k}$
(b) $4x + 10y + z = 21$, $\mathbf{r} = (2 + 4t)\mathbf{i} + (1 + 10t)\mathbf{j} + (3 + t)\mathbf{k}$] □

Suppose now that we want to calculate the rate of change of $V(\mathbf{r})$ in a particular direction specified by the unit vector $\mathbf{t}$. Let s be the distance travelled in the direction of $\mathbf{t}$; then $\mathrm{d}\mathbf{r} = \mathbf{t}\,\mathrm{d}s$. So $\mathrm{d}V = \nabla V.\mathbf{t}\,\mathrm{d}s$. Hence we can conclude that the rate of change of V in the direction of $\mathbf{t}$ is

$$\frac{\mathrm{d}V}{\mathrm{d}s} = \nabla V.\mathbf{t} = \mathbf{t}.\nabla V.$$

$\mathbf{t}.\nabla V$ is called the *directional derivative*. Now

$$\nabla V.\mathbf{t} = |\nabla V|\,|\mathbf{t}| \cos\theta = |\nabla V| \cos\theta,$$

where θ is the angle between the vectors ∇V and $\mathbf{t}$. This is maximized when $\cos\theta = 1$, i.e. when $\theta = 0$. Thus V changes most rapidly in the direction of ∇V, and $|\nabla V|$ is this most rapid rate of change. It is this property,

in the two-dimensional case, that gave rise to the name gradient, because $|\nabla V|$ is the gradient of the surface given by $z = f(x, y)$ in that case. Correspondingly the maximum decrease is when **t** is opposite to ∇V.

Example 1.8. Find the directions in which the function $f(x,y,z) = (x/y) - yz$ increases and decreases most rapidly at the point P, $(4,1,1)$.

We can describe the directions in which f increases and decreases most rapidly by specifying the unit vectors in those directions. Now

$$\nabla f = \frac{1}{y}\mathbf{i} - \left(\frac{x}{y^2} + z\right)\mathbf{j} - y\mathbf{k} = (1, -5, -1) \text{ at P.}$$

The rate of change of f in the direction of unit vector **t** is $\nabla f.\mathbf{t}$. This has its maximum when **t** is in the same direction as ∇f; so the direction **t** in which f increases most rapidly is

$$\frac{\nabla f}{|\nabla f|} = \frac{1}{\sqrt{27}}(1, -5, -1).$$

and the actual rate is $\sqrt{27}$. The rate of *decrease* of f is greatest, at $-\sqrt{27}$, when **t** is in the *opposite* direction, i.e.

$$\frac{-1}{\sqrt{27}}(1, -5, -1)$$

Exercise 1.6. Find the directional derivative of Φ at the point $(1, 2, 3)$ in the direction of the vector $(1, 1, 1)$ where

$$\Phi = \frac{x^2}{3} + \frac{y^2}{9} + \frac{z^2}{27}.$$

□

We can write ∇ on its own as

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$$

and work with it like a vector field, although it is in fact not a vector field (since we cannot say what numerical value its components have at a particular point); strictly speaking ∇ is a **vector differential operator**. The name of the symbol ∇ is 'nabla' but often in speech we say 'del'. It is easy to see how to take a two-dimensional version of ∇. We will return to ∇ in Chapter 3, where we shall see how the ∇ operator can also be used to differentiate vectors.

Chapter 2

Curves, Lengths, Surfaces and Areas

Last revised 13 Oct 2010.

This chapter gives an overview of various methods for describing curves in 2-dimensional and 3-dimensional space, including parametrised curves, curves in polar coordinates, and some vector notation.

We then move on to formulae involving integration for the arc-length of curves in each of these cases, and for the area enclosed by curves in 2-dimensional polar coordinates.

Finally we extend this to defining parametrised surfaces in 3 dimensions (using 2 parameters), and the areas of these surfaces.

2.1 Parametrised curves

2.1.1 Parametrised curves: definition

(See Thomas 3.5)

By now you are familiar with expressing a curve on a 2-dimensional plane in Cartesian coordinates (x,y) as

$$y = f(x)\ ; \tag{2.1}$$

where $f(x)$ is a given function, and may be any combination of polynomials, trigonometric functions, exponential functions (called "elementary functions"), or more complicated functions. This form for a curve is called "explicit form" since a given f specifies exactly how to calculate y for any value of x. Clearly for a given function $f(x)$ we can draw a graph of this function by taking many values of x with suitably small steps, evaluating $f(x)$ at each of these so we have a "dot" at $(x,y = f(x))$ and then "joining the dots". If f is a continuous function, then there are no "gaps" in this curve.

This is straightforward, but we have the limitation that for each x the curve has a unique value of y (the converse is not true, i.e. choosing a particular value y_0 for y and solving the equation $f(x) = y_0$ may give none, one or many solutions known as "roots" for x) . So, a curve such as $y = f(x)$ can have "wiggles" in the y-direction but not in the x-direction.

In a few special cases we can have multiple values for y at a given x, e.g. for the familar case of the circle $x^2+y^2=a^2$ we can write $y=\pm\sqrt{a^2-x^2}$, and the $\pm$ term gives 2, 1 or 0 solutions for y depending whether $|x|<a$, $x=a$, or $|x|>|a|$ respectively ; but this quickly becomes excessively complicated for more general curves.

A second way to represent a curve in a plane is as the set of all points satisfying an equation such as $f(x,y)=0$ or $f(x,y)=c$, where c is a constant; here f depends on both x and y and may not be separable; this is called "implicit form". This has some advantages we will see later, for example choosing different values of c can give us a "family" of different curves from one function f; however, a clear disadvantage is that there is no easy way (in general) to calculate y at a given $x=x_0$; so sketching the curve (or programming a computer to sketch it) can be cumbersome, unless we can recognise the form of the solution from experience.

A third way to represent a curve in a 2-dimensional plane is via **parametrisation**: now we define **two** arbitrary functions $f(t),g(t)$ of a new real variable t, and we define our curve called C as the set of all points where

$$x=f(t),\quad y=g(t)\qquad \text{hence}\quad (x,y)=(f(t),g(t)) \tag{2.2}$$

for any value of t in a given **domain** (which may be finite or infinite). Here we call C the **parametric curve**, t is the **parameter**, and $x=f(t),y=g(t)$ are the **parametric equations** for the curve. These equations together with the defined domain of t constitute a complete definition called the **parametrisation** of the curve.

Given the above, it is clear that any value of t maps to a single point in the (x,y) plane; it is also fairly obvious that if the functions f,g are both continuous, then the resulting curve C is also continuous. (To prove this, pick a value t_0 giving a point on the curve x_0,y_0; then draw a tiny circle radius δ around x_0,y_0. If f,g are both continuous we can find some range of $t\pm\varepsilon$ for which the curve is contained inside the above circle, i.e. the curve has no "gaps"; if there were a finite gap in the curve, then either f or g must not be continuous, contradicting our assumption).

Using parametric form, we can express more complicated curves such as figure-eights, spirals and so on which can self-intersect and/or cross a given x value many times (including infinitely many), and we can "sketch the curve" by hand or by computer by just evaluating $f(t),g(t)$ at a sufficient number of points spaced in t and "joining the dots".

Note that here t is not necessarily "time", t is just a "label" so that each point on the curve is "labelled" with one value of t, or multiple values if the curve crosses itself at that point.

Clearly if we are given a curve $y=g(x)$, we can put that into parametric form by simply defining $f(t)=t$ in the above, so then $x=t$ and $y=g(t)=g(x)$; but the converse generally is not true, so the parametric form is more general.

Now for a few simple examples: a very simple example is a straight line, which is given by

$$x=x_0+at,\qquad y=y_0+bt\quad ; \tag{2.3}$$

and the domain $-\infty<t<\infty$. Here it is easy to see that this parametrises a straight line passing through the point (x_0,y_0) with direction vector (a,b) and slope b/a. (If $a\neq 0$, we can rearrange the $x-$equation to $t=(x-x_0)/a$, and then substitute that into the $y-$ equation to get $y=y_0+(b/a)(x-x_0)$.)

Note that many possible choices of x_0,y_0,a,b lead to the same straight line, only the mapping from t onto points on the line will change. If we want a line through (x_0,y_0) and (x_1,y_1) then we set $a=x_1-x_0,\ b=y_1-y_0$ in the above, and if we want our "curve" to be a finite straight line segment with endpoints (x_0,y_0) and (x_1,y_1), then we just specify that the domain of t is $0\le t\le 1$ above.

Another simple case is given by

$$x=a\cos t, y=a\sin t\quad ;$$

it is clear that this obeys $x^2+y^2=a^2$, so the curve C is a circle of radius a centred at the origin. Since the functions sin and cos are both periodic with period 2π, adding $n\times 2\pi$ to t (for any integer n) gets back to the same x,y. So if we let t run from $-\infty$ to $+\infty$, the resulting curve loops around the circle an infinite number of times. So, in this case it is more convenient to specify a finite domain for t, such as $0\le t<2\pi$, so over that range the curve goes round the circle exactly once. (Here we can choose any interval of length 2π, so e.g. $-\pi<t\le\pi$ works just as well).

We can generalise this to an ellipse by $x=a\cos t, y=b\sin t$; this is just a circle stretched by a factor b/a in the $y-$direction, so the semi-axes are a and b respectively. We can get an ellipse with mid-point at (x_0,y_0) with by $x=x_0+a\cos t, y=y_0+b\sin t$.

The above illustrates a convenient property: because if we know one parametric curve, we can produce a shifted copy of it just by adding x_0 and y_0 to the two functions; or we can stretch or squash it along the axes by multiplying our two functions by constants.

2.1.2 The cycloid

Another example of a curve which is easy to represent in parametric form is the **cycloid**, which can be expressed as

$$x=a(t-\sin t)\ , \qquad y=a(1-\cos t) \tag{2.4}$$

If we didn't have the at term in the $x-$equation above, it is easy to see this would be a circle of radius a centred at $(0,a)$; but the additional at term makes the circle's centre "roll along" in the x direction as t increases. It turns out that the above curve is the curve traced out by (for example) a pebble stuck to a bicycle's tyre as the tyre rolls along the ground without slipping, so we get a combination of the "axle" going along at constant rate and the pebble going in a circle round the moving axle. In the example above we have chosen things so the "ground" is the x-axis, the axle goes along the line $y=a$ and the point is at the origin at $t=0$.

This curve has applications in several real-world problems, and you can see above that it is quite simple to write in parametric form, but it is complicated in Cartesian coordinates (there is an expression in elementary functions for x in terms of y, but not the other way round).

There are generalised versions of this curve called the epicycloid and hypocycloid which are traced by a point on one circle rolling around a second circle (instead of along a straight line), and furthermore there are versions where the "point" is not on the circumference of the rolling circle; these are called **trochoids**. (You won't be expected to memorise these, but you might be given the equations as part of an exam question so it is worth knowing the general concept).

2.1.3 Lissajous figures

A curve parametrised by $x=a\cos k_1 t,\quad y=b\sin k_2 t$ where k_1,k_2 are constants (usually integers), is called a **Lissajous figure**. By considering what happens as t varies, we can see that both x and y oscillate between $\pm a$ and $\pm b$, so the curve must always lie inside a rectangle with corners at $(\pm a,\pm b)$; but now the curve oscillates at different rates in the x,y directions, and it can cross itself many times. If we choose $k_1=1,k_2=2$ it will turn out that we get a figure-of-eight. If k_1/k_2 is a simple fraction, it will turn out that the curve closes back on itself after a finite number of "wiggles" ; but if k_1/k_2 is irrational it can be shown that the curve gets arbitrarily close to every point in the above rectangle but never returns to exactly the same place; this sort of thing may be seen in some computer screensavers, where you have an icon wandering around the rectangular computer screen and it's helpful for the pattern not to repeat itself.

2.1.4 Parametric curves in 3 dimensions

We can easily extend the above parametric curves from 2 to 3 dimensions by defining a third function $h(t)$ for the $z-$coordinate, so that

$$x = f(t),\ y = g(t),\ z = h(t). \tag{2.5}$$

Clearly for each value of t we now get a point in 3-dimensional space, and the set of points $(f(t), g(t), h(t))$ defines a 1-dimensional curve which is continuous if f, g, h are all continuous; the basic principles are the same as in 2 dimensions.

A good example of this is the **helix**, where

$$x = a\cos t\ , \qquad y = a\sin t\ , \qquad z = bt \tag{2.6}$$

where a, b are constants. Here as t varies, the distance of the curve from the z-axis is $\sqrt{x^2+y^2} = a$ (constant), so the curve projected onto the x,y plane is a circle, but the z-value is increasing at a uniform rate, so we get a curve in 3 dimensions looking like the handrail of a spiral staircase winding around the z-axis. Each increase of 2π in t gives us one full "twist" around the $z-$axis.
Note: In everyday English, this may be called a spiral: however in maths terminology, the term spiral refers to various types of 2-dimensional plane curve, while a 3-dimensional curve as described above is properly called a helix).

Parametric curves may be expressed more compactly in vector notation as $\mathbf{r} = \mathbf{r}(t)$, but of course we still need to define the 3 functions for the 3 independent components of $\mathbf{r}$, so this doesn't change any of the results, it just makes the expressions more compact.

As an example, the parametric representation also makes it quite easy to express curves which aren't symmetric about the x,y,z axes: for example, if we choose any two fixed orthogonal unit vectors $\mathbf{u}, \mathbf{v}$, we can construct an ellipse with centroid at $\mathbf{c}$, semi-major axis a and semi-minor axis b respectively parallel to the two vectors $\mathbf{u}, \mathbf{v}$, by:

$$\mathbf{r}(t) = \mathbf{c} + (a\cos t)\mathbf{u} + (b\sin t)\mathbf{v} \quad ; \tag{2.7}$$

we can of course plug in the components to write x,y,z in terms of t, but then it will be a lot less clear geometrically.

2.2 Arc Length of a curve

Here we show how to calculate the arc-length of a curve between two given endpoints.

If we choose a point on the curve $\mathbf{r}(t)$, and a neighbouring point $\mathbf{r}(t+\delta t)$, then the vector difference of these is

$$\mathbf{r}(t+\delta t) - \mathbf{r}(t) \approx \frac{d\mathbf{r}}{dt}\,\delta t \quad ; \tag{2.8}$$

this is the vector separation between the two nearby points on the curve. Taking limits where δt tends to zero, and assuming that the derivative exists, the curve tends to an infinitesimal straight line segment, so we can define the infinitesimal **length** ds to be the modulus of the left-hand side above,

$$ds = |\mathbf{r}(t+dt) - \mathbf{r}(t)| = \left|\frac{d\mathbf{r}}{dt}\right| dt \tag{2.9}$$

$$= \left|\left(\frac{df}{dt}, \frac{dg}{dt}, \frac{dh}{dt}\right)\right| dt \tag{2.10}$$

Therefore, we have

$$\frac{ds}{dt} = \sqrt{\left(\frac{df}{dt}\right)^2 + \left(\frac{dg}{dt}\right)^2 + \left(\frac{dh}{dt}\right)^2} \tag{2.11}$$

$$\text{or } \frac{ds}{dt} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} \tag{2.12}$$

(This is effectively just Pythagoras's theorem applied to an infinitesimal segment of the curve, which has the same length as a straight line joining its endpoints).

So, now we can define the **arc length** L of the parametric curve between two values t_1, t_2 by integrating the above, giving us

$$L = \int_{t_1}^{t_2} \sqrt{\left(\frac{df}{dt}\right)^2 + \left(\frac{dg}{dt}\right)^2 + \left(\frac{dh}{dt}\right)^2}\, dt \tag{2.13}$$

To be clear, L is the length of a virtual "piece of string" which exactly follows the curve between endpoints given by t_1 and t_2 at points $(f(t_1), g(t_1), h(t_1))$ and $(f(t_2), g(t_2), h(t_2))$, if the string was then "pulled out straight". The above L is **not** the straight-line distance between the endpoints which would be just $|\mathbf{r}(t_2) - \mathbf{r}(t_1)|$. Equation 2.13 remains valid even if some or all of the derivatives cross zero, as long as none of them become infinite or undefined. If our curve is in 2 dimensions x, y we just set $z = h(t) = 0$ and $dz/dt = 0$.

Note: In problems, you may be given a parametric curve, and the endpoints specified in terms of (x_1, y_1, z_1) and (x_2, y_2, z_2); in this case you will need to solve to find the values of t_1 and t_2 corresponding to the endpoints, before doing the integral above. For each endpoint you can solve whichever of the x, y, z equations is simplest to get t_1, t_2; then insert those t_1, t_2 into the other two equations to check.

Example 2.1. The parametric curve C is given by $x = t, y = t^2, z = \frac{2}{3}t^3$. Evaluate the arc-length L of the curve between points (0,0,0) and $(2, 4, \frac{16}{3})$.

Answer: The end-points have values $t_1 = 0$ and $t_2 = 2$ (solve the x equation for t, and check the other two equations give the desired point); the derivatives are $dx/dt = 1$, $dy/dt = 2t$, $dz/dt = 2t^2$. Therefore the required length is

$$L = \int_{t=0}^{2} \sqrt{1 + (2t)^2 + (2t^2)^2}\, dt = \int_0^2 \sqrt{1 + 4t^2 + 4t^4}\, dt = \int_0^2 1 + 2t^2\, dt = [t + \tfrac{2}{3}t^3]_0^2 = \tfrac{22}{3} \quad .$$

Equation 2.13 can easily be simplified to give us the arc-length of a curve in implicit form: i.e. if we are given a 2-dimensional curve given as $y = g(x)$, we can just define $f(t) = t$ so $t = x$, $y = g(t) = g(x)$ and $z = h(t) \equiv 0$; inserting this gives us the arc-length for the curve $y = g(x)$ between the endpoints at $(x_1, g(x_1))$ and $(x_2, g(x_2))$ as

$$L = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \tag{2.14}$$

Similarly, if we have a curve in 3 dimensions where any two of the coordinates are given as functions of the other one, e.g. $y = g(x), z = h(x)$, then we get

$$L = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2 + \left(\frac{dz}{dx}\right)^2}\, dx \quad . \tag{2.15}$$

for L the arc-length between end-points at x_1 and x_2.

(However, note that many common curves have arc-length integrals which are not soluble in elementary functions; an ellipse is a well-known case where the circumference is not elementary, but is given by a special function called an *elliptic integral*. But one-dimensional integrals are generally very fast to evaluate accurately with a computer, since only one loop is needed).

2.2.1 Tangent vector to a curve

Given a parametric curve in either 2 or 3 dimensions, we can clearly differentiate each of the component functions with respect to t ; assuming the functions are differentiable, this gives us a new vector

$$\frac{d\mathbf{r}}{dt} = \lim_{\delta t \to 0} \frac{\mathbf{r}(t+\delta t) - \mathbf{r}(t)}{\delta t} = \left(\frac{df}{dt}, \frac{dg}{dt}, \frac{dh}{dt} \right) \quad . \tag{2.16}$$

It is easy to see that this vector $d\mathbf{r}/dt$ is locally parallel to the curve at the selected point, as long as all the derivatives exist ; hence given a value of $t = t_0$, we can evaluate $\mathbf{r}_0 = \mathbf{r}(t_0)$ (which is the position vector of a point on the curve), and define $\mathbf{Q} = d\mathbf{r}/dt|_{t_0}$ which is a vector of direction tangent to the curve at the same point, so $\mathbf{Q}$ is a **tangent vector** to the parametric curve at the point $\mathbf{r}_0$. Thus, we can construct an equation for the **tangent line**

$$\mathbf{r} = \mathbf{r_0} + u\mathbf{Q} \quad , \quad -\infty < u < \infty \tag{2.17}$$

for any real u; here u is another parameter (giving position along the tangent line to our curve C at $\mathbf{r}_0$). If we write this out in components, the above is three linear equations giving x, y, z as linear functions of u, and if desired we can rearrange those to give two linear equations for the tangent line: e.g. if we want y and z in terms of x, we just rearrange the x equation to give u in terms of x, and substitute that into the y, z equations.

Note: in the above, we must evaluate $\mathbf{r}$ and $d\mathbf{r}/dt$ at the same point i.e. the same value of t, otherwise the result will not make sense. Also, if you are given the coordinates of $\mathbf{r}_0$ rather than t, you will have to find the value of t which gives you $\mathbf{r}(t) = \mathbf{r}_0$; you can pick whichever coordinate is the simplest to solve.

(**Warning:** equation 2.16 looks a bit like the equation for ∇f which we met earlier. However, it's actually very different because $f(\mathbf{r})$ was a scalar function of three variables x, y, z, while $\mathbf{r}(t)$ along a parametric curve is a vector-valued function of one variable t.)

2.3 Curves in polar coordinates

As we saw in the previous section, it is sometimes convenient if we are working with circles, ellipses or other closed curves to work in **plane polar coordinates**; here instead of the familiar x, y of Cartesian coordinates, we can label any point P in a plane by its distance r from a fixed origin O, and an angle θ between the line OP and the positive $x-$axis. By convention, θ is defined to increase "anticlockwise" (from $+x$ towards $+y$) so the positive $y-$axis has $\theta = +\frac{\pi}{2}$, the negative $x-$axis has $\theta = \pi$, and the negative $y-$axis has $\theta = \frac{3\pi}{2}$.

The conversion from (r, θ) to x, y is given by simple trigonometry:

$$x = r\cos\theta \qquad y = r\sin\theta \quad , \tag{2.18}$$

which is clearly unique. Rearranging these to give r, θ in terms of x, y, we have

$$r = \pm\sqrt{x^2 + y^2} \qquad \theta = \arctan(y/x) + (n\pi) \tag{2.19}$$

for some integer n. Note however that the conversion from (x, y) to (r, θ) is *non-unique*; one point in a plane has a unique (x, y) pair, and a particular pair (r, θ) map to a unique x, y; but one point (x, y) can be represented

by two possible values of r (positive and negative) and an infinite number of θ values differing by integer multiples of π i.e. different numbers of half-turns around the origin. So the point (r,θ) is actually the same point as $(-r,\,\theta+(2n+1)\pi)$ and $(+r,\,\theta+2n\pi)$.

In most cases of interest we will be taking the positive value of r and we'll take θ to lie in the interval $[0,2\pi]$, in which case the mapping is unique, but be aware of potential ambiguities with this.

Clearly one of the simplest curves in plane polar coordinates is

$$r = a \tag{2.20}$$

where a is a constant. Implicitly this also means $\theta =$ any value, so this is clearly a circle, centred on the origin, radius a. Likewise $\theta = b$ where b is a constant is a straight line through the origin at angle b.

More generally we can define curves as

$$r = f(\theta) \tag{2.21}$$

This is a convenient way to define certain types of curve; one example is

$$r = a + b\theta \tag{2.22}$$

which describes a spiral called an **Archimedes spiral**.

It is possible to express a straight line in polars, for example it is easy to show that the vertical line $x = b$ has the polar equation $r = b\sec\theta$. More generally for a line which has closest distance b to the origin at angle θ_0, we get $r = b\sec(\theta-\theta_0)$.

2.3.1 Conics in polar coordinates

A particularly useful family of curves in polar coordinates is given by

$$r(\theta) = \frac{\ell}{1+e\cos\theta} \tag{2.23}$$

where e is called the eccentricity and ℓ is the semi-latus rectum. It can be shown that this equation gives a **conic section** as we met before with the quadratic functions in x,y. Here $e = 0$ gives a circle (with radius ℓ), $0 < e < 1$ gives an ellipse, $e = 1$ gives a parabola and $e > 1$ gives a hyperbola, so by choice of e this expression can give any of the above conics ; and ℓ is just a scale factor giving the overall size.

Note: in the above representation Eq 2.23, the origin is one *focus* of the conic, the origin is not the centroid except if $e = 0$ (the circle). The form Eq. 2.23 is especially useful in astronomical orbit problems, since it will turn out that orbits of planets and comets around a central star have a solution of this type with the star at one focus (and nothing at the other focus). For ellipses, it is easy to show by plugging in $\theta = 0,\pi$ that the semi-major axis $a = \ell/(1-e^2)$. (Note we use ℓ rather than a in the above definition since ℓ is well-defined for all of the conics).

For any point given by eq 2.23, the distance to the origin is r and the distance d to a vertical line at $x = x_0$ is

$$d = x_0 - r\cos\theta = \frac{x_0 + x_0 e\cos\theta - \ell\cos\theta}{1+e\cos\theta} \qquad . \tag{2.24}$$

If we choose $x_0 = k = \ell/e$, this reduces to $d = r/e$ or $r = ed$, so our conic is the locus such that (distance from focus) = $e\times$ distance from the line $x = k$, called the **directrix** ; this works for any of the ellipse, parabola or hyperbola; though the directrix is "at infinity" for the circle $e = 0$.

Altogether there are (at least) four possible ways of defining the conic sections: one is the "plane slicing a cone" definition; the second is using the distances-from-foci properties; the third is via quadratic equations in Cartesian coordinates, and the fourth is as above. (It takes some straightforward but fairly long algebra to prove that all of these do actually end up with the same family of curves, which we won't repeat here (see e.g. Thomas Chapter 10)).

2.3.2 Arc length and area in plane polar coordinates

Given a curve in polar coordinates as $r = f(\theta)$, we can get the arc length in two ways: firstly we can put this into the parametric representation by $x = f(\theta)\cos\theta, y = f(\theta)\sin\theta$ where θ is the parameter (which behaves like t in the examples we saw before). If we differentiate the above, we have

$$\frac{dx}{d\theta} = -f(\theta)\sin\theta + \frac{df}{d\theta}\cos\theta \qquad \frac{dy}{d\theta} = f(\theta)\cos\theta + \frac{df}{d\theta}\sin\theta \tag{2.25}$$

Inserting these into equation 2.11 for the arc-length, we get

$$\frac{ds}{d\theta} = \sqrt{(f(\theta))^2 + \left(\frac{df}{d\theta}\right)^2}\,, \tag{2.26}$$

therefore the arc-length of the curve defined by $r = f(\theta)$ between endpoints given by $\theta = \theta_1$ and $\theta = \theta_2$ is

$$L = \int_{\theta_1}^{\theta_2} \sqrt{(f(\theta))^2 + \left(\frac{df}{d\theta}\right)^2}\, d\theta \tag{2.27}$$

(We can get the same result geometrically by drawing a segment of a curve from (r,θ) to $(r+\delta r, \theta+\delta\theta)$, also drawing the circular arc through (r,θ), and applying Pythagoras's theorem to the small triangle resulting).

Warning: the above is clearly different from Equation 2.14 which gave the arc-length for the case $y = g(x)$; comparing them, the second term looks the same, but the first term above is $f(\theta)^2$ instead of 1. The reason is that in polar coordinates, a small change of angle $\delta\theta$ shifts our point by a distance $r\,\delta\theta$ in the "circumferential" (around-the-origin) direction, **not** just $\delta\theta$. We will see a lot more of this sort of thing in later sections where we deal with 3-dimensional polar coordinates.

2.3.3 Area in polar coordinates

If we are given a curve $r = f(\theta)$, it is straightforward to evaluate the area of the sector bounded by two straight lines $\theta = a$, $\theta = b$ and the curve $r = f(\theta)$: by considering an interval from $(f(\theta_0),\theta_0)$ to $(f(\theta_0+\delta\theta),\theta_0+\delta\theta)$, if $\delta\theta$ is small this area approaches an isosceles triangle with long dimension $r = f(\theta)$ and width $r\,\delta\theta$, so the area is $\frac{1}{2}r^2\delta\theta$.

Thus, the area inside a curve defined in plane polar coordinates, between angles $\theta_1 \le \theta \le \theta_2$ is simply

$$A = \frac{1}{2}\int_{\theta_1}^{\theta_2} [r(\theta)]^2\, d\theta \qquad . \tag{2.28}$$

Warning: You need to beware of zero-crossings here: if $r(\theta)$ goes negative so the curve has several "petals", you need to be careful not to count the same petal twice at angles π apart. If this happens, it's advisable to sketch the curve, break the integral into suitable chunks where r does not cross zero, and add these up. If r is always non-negative, there are no problems.

Example 2.2. An easy example is the circle $r = a$: inserting $\theta_1 = 0, \theta_2 = 2\pi$ for the endpoints (as explained earlier) gives us

$$A = \tfrac{1}{2}\int_0^{2\pi} a^2 d\theta = \pi a^2. \tag{2.29}$$

giving us the familiar formula for the area of a circle.

Example 2.3. Another example is the **cardioid** defined by $r = a(1+\cos\theta)$; this curve has real-world applications since many microphones and radio antennas have a directional response given by a cardioid function. Inserting this into 2.28, and using the double-angle formula we easily get the area $\frac{3}{2}\pi a^2$.

2.4 Surfaces in 3 dimensions

In the previous section we looked at parametric representations of a curve in 2 or 3 dimensions, generally written as $(x,y,z) = (f(t),g(t),h(t))$; there we had 2 or 3 functions (one per coordinate) of **one** parameter t.

A further generalisation is to define a **surface** in 3-dimensional space; a plane is the simplest example, but in general we will deal with curved surfaces. We will see that it requires **two** parameters to describe an arbitrary surface, instead of the one parameter we had for a curve.

Now we'll call the parameters u and v, and as before we need 3 functions of these to define the 3 x,y,z coordinates of our surface in 3-dimensional space, so we get a **parametrised surface** as

$$x = f(u,v),\ y = g(u,v),\ z = h(u,v) \quad \text{i.e.} \quad \mathbf{r} = \mathbf{r}(u,v)\ . \tag{2.30}$$

If we pick a fixed value for v, say $v = v_0$, and allow u to vary, then we just have a one-dimensional curve (on the surface) as we vary u. If we now choose $v = v_0 + \Delta v$ and vary u again, we get another curve which is "close" to the first one if the functions are continuous, and we have a "ribbon" of surface bounded by the two curves. Repeating for lots of v_0's we see that we sweep out a 2-dimensional surface (call it S) in 3-dimensional space, as long as the curves for $v = v_0$ and $v = v_0 + \Delta v$ don't coincide. Technically we can define the partial derivatives

$$\frac{\partial \mathbf{r}}{\partial u} = \left(\frac{\partial f}{\partial u}, \frac{\partial g}{\partial u}, \frac{\partial h}{\partial u}\right)\ , \qquad \frac{\partial \mathbf{r}}{\partial v} = \left(\frac{\partial f}{\partial v}, \frac{\partial g}{\partial v}, \frac{\partial h}{\partial v}\right) \tag{2.31}$$

and as long as these two vectors are not parallel at any point, our locus of $\mathbf{r}(u,v)$ will in fact be a surface, not a line. We can see that both of the above two vectors are directions tangent to the surface at $\mathbf{r}(u,v)$. We can also take the vector product of these two,

$$\mathbf{N} = \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \qquad . \tag{2.32}$$

This cross-product $\mathbf{N}$ will be non-zero if the two partial derivatives above are both non-zero and not parallel. Assuming $\mathbf{N}$ is non-zero, it must be a **normal vector** to the surface S, because both the partial derivatives are parallel to the tangent plane to our surface at the point $\mathbf{r}(u,v)$, and $\mathbf{N}$ is perpendicular to both of them.

Thus, if we are given a surface $\mathbf{r}(u,v)$, and given a point on the surface defined by values (u_0,v_0), we have a clear procedure for finding the tangent plane to the surface at the corresponding point: we first evaluate the point in the surface $\mathbf{r}_0 = \mathbf{r}(u_0,v_0)$; next we evaluate the two partial derivatives at the same point, and take their cross-product $\mathbf{N}$ as above; thus in the usual vector notation for a plane through a given point normal to a given vector, the equation for the tangent plane is $(\mathbf{r} - \mathbf{r}_0).\mathbf{N} = 0$.

(Note: if instead we are given a point in the surface by defining its (x,y,z) space coordinates, we will first have to find the values of (u_0,v_0) which map onto that point *before* we evaluate the partial derivatives; in general that may not be simple to do, but it usually will be in the case of test questions.)

We can also use vectors to calculate the area of a parametric surface given by $\mathbf{r}(u,v)$: if we take the four points $\mathbf{r}(u,v)$, $\mathbf{r}(u+du,v)$, $\mathbf{r}(u,v+dv)$, $\mathbf{r}(u+du,v+dv)$, these define an infinitesimal parallelogram with sides $\frac{\partial \mathbf{r}}{\partial u}du$ and $\frac{\partial \mathbf{r}}{\partial v}dv$; as we saw from the definition of the vector product in Chapter 1, the area dA of this parallelogram is

$$dA = \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| \, du\, dv \tag{2.33}$$

Thus, integrating the above with respect to both of u,v we get the **surface area** A of our parametric surface as

$$A = \int\int_D \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| \, du\, dv \tag{2.34}$$

where the domain D of integration is the appropriate domain of u,v. (Note: if the surface is described geometrically, we will need to work out limits on u,v to cover the described surface).

In the special case where our surface is given as $z = h(x,y)$, we can just substitute $x = u$, $y = v$ into the above: then the two partial derivative vectors become $(1,0,\partial h/\partial x)$ and $(0,1,\partial h/\partial y)$, and the surface area becomes

$$A = \int\int_D \sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2 + 1}\, dx\, dy \tag{2.35}$$

where the integral is over some given domain in x,y.

Example 2.4. A good example of the above is the area of a sphere: a parametrisation of a sphere in 2 parameters (θ,ϕ) is

$$x = a\sin\theta\cos\phi, \quad y = a\sin\theta\sin\phi, \quad z = a\cos\theta \qquad 0 < \theta < \pi,\ 0 < \phi < 2\pi \tag{2.36}$$

(It is easy to show that this satisfies $x^2+y^2+z^2 = a^2$, so any (x,y,z) above does lie on the sphere. I'll state without proof that the limits given above define a unique mapping from a point on the sphere to θ,ϕ. We will meet this again later when we come to spherical polar coordinates).

Given this, evaluating the partial derivatives, we have $\partial\mathbf{r}/\partial\theta = (a\cos\theta\cos\phi, a\cos\theta\sin\phi, a\sin\theta)$, and $\partial\mathbf{r}/\partial\phi = (-a\sin\theta\sin\phi, a\sin\theta\cos\phi, 0)$. The cross product of these vectors is
$\mathbf{N} = (a^2\sin^2\theta\cos\phi, a^2\sin^2\theta\sin\phi, a^2\sin\theta\cos\theta)$ which is $a\sin\theta\mathbf{r}$, and has magnitude $a^2\sin\theta$. Then our surface area A of the sphere becomes

$$\begin{aligned} A &= \int_0^{2\pi}\left(\int_0^{\pi} a^2\sin\theta\, d\theta\right) d\phi \\ &= \int_0^{2\pi} [-a^2\cos\theta]_0^{\pi}\, d\phi \\ &= \int_0^{2\pi} 2a^2\, d\phi \\ &= 4\pi a^2 \qquad . \end{aligned}$$

2.4.1 Parametric forms of common surfaces

To conclude this chapter, I'll give some specific examples of parametric forms for common surfaces. These will turn out to be useful later, when we come to evaluate integrals over specified 2D surfaces in 3D space: the parametric form is usually the easiest way to do this.

A plane in 3D can be expressed in parametric form as

$$\mathbf{r}(u,v) = \mathbf{r}_0 + u\mathbf{a} + v\mathbf{b}$$

where $\mathbf{r}_0$ is a point in the plane, and $\mathbf{a}$, $\mathbf{b}$ are any two vectors parallel to the plane. Here if we take $-\infty < u,v < \infty$ we get the whole infinite plane.

If we want a finite parallelogram with one corner at $\mathbf{r}_0$ and two adjacent sides $\mathbf{a}$, $\mathbf{b}$, we can simply put limits $0 \le u \le 1,\ 0 \le v \le 1$ in the above. (A rectangle is a special case of this).

Finally, I'll repeat the parametric forms for a cylinder, sphere, ellipsoid and hyperboloid which we met briefly in Chapter 1.

	Implicit form	Parametric form		
Cylinder	$x^2+y^2=a^2$	$(x,y,z) = (a\cos\theta, a\sin\theta, z)$	Parameters θ, z	(2.37)
Sphere	$x^2+y^2+z^2=a^2$	$(a\sin\theta\cos\phi, a\sin\theta\sin\phi, a\cos\theta)$	Parameters θ, ϕ	(2.38)
Ellipsoid	$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=+1$	$(a\sin\theta\cos\phi, b\sin\theta\sin\phi, c\cos\theta)$	Parameters θ, ϕ	(2.39)
Hyperboloid	$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$	$(a\cos u, b\sin u\cosh v, c\sin u\sinh v)$	Parameters u, v	(2.40)

In the above, the choice of the names for the two parameters is slightly arbitrary, but follows common conventions.

It is easy to show that each parametric form above satisfies the implicit-form equation, just by substituting and using $\sin^2 A + \cos^2 A = 1$ or $\cosh^2 A - \sinh^2 A = 1$. It's not so obvious to see how to go the other way; but the parametric form for the sphere falls out naturally when we come to spherical polar coordinates; the ellipsoid is a simple "stretch" of the sphere; and the hyperboloid comes from replacing sin with sinh etc in the ellipsoid.

Chapter 3

Vector differentiation, the ∇ operator, grad, div and curl.

Last update: 20 Oct 2010.

Syllabus topics covered:
1. Vector fields
2. Grad, div and curl operators in Cartesian coordinates. Grad, div, and curl of products etc.

Here we cover differentiation of vectors. Note that this differs from the gradient introduced in Chapter 1, where we obtained a vector by differentiating a scalar field.

3.1 Vector functions of one or more variables

(See Thomas 13.1)
In many physical contexts one is interested in vectors that vary with position or time. For example, the position of a point can be described by a vector $\mathbf{r}$. Thus, if we consider a moving particle, its position can be described as a function of time t by the vector $\mathbf{r}(t)$, and its rate of change with respect to t will be the velocity (which has magnitude and direction, i.e. is a vector: its magnitude is the speed). The position vector is then a function of one variable.

Another context is where we have a vector defined at each point, say $\mathbf{F}(\mathbf{r}) = \mathbf{F}(x, y, z)$ and a curve with a parameter u, say, so its points are $(x(u), y(u), z(u))$. Then we can define a vector function of u, $\mathbf{F}(u) = \mathbf{F}((x(u), y(u), z(u))$. We can deal with this and the moving particle case as follows.

A vector function of a scalar u, $\mathbf{F}(u)$, can be defined by specifying its components as functions of u:

$$\mathbf{F}(u) = (f_1(u), f_2(u), f_3(u)) .$$

The derivative $\mathrm{d}\mathbf{F}/\mathrm{d}u$ of $\mathbf{F}$ with respect to u is then:

$$\frac{\mathrm{d}\mathbf{F}}{\mathrm{d}u} = \left(\frac{\mathrm{d}f_1}{\mathrm{d}u}, \frac{\mathrm{d}f_2}{\mathrm{d}u}, \frac{\mathrm{d}f_3}{\mathrm{d}u} \right) .$$

This simply goes back to the fundamental definition of a derivative:

$$\frac{\mathrm{d}\mathbf{F}}{\mathrm{d}u} = \lim_{\delta u \to 0} \frac{\mathbf{F}(u+\delta u) - \mathbf{F}(u)}{\delta u}.$$

Clearly one can compute higher derivatives, such as $\mathrm{d}^2\mathbf{F}/\mathrm{d}u^2$, by differentiating the components of $\mathbf{F}$ the required number of times.

Example 3.1. If $\mathbf{r}(t)$ is the position vector of a particle, as a function of time t, then $\mathrm{d}\mathbf{r}/\mathrm{d}t$ is the velocity $\mathbf{v}$ of the particle. Also $\mathrm{d}\mathbf{v}/\mathrm{d}t \equiv \mathrm{d}^2\mathbf{r}/\mathrm{d}t^2$ is the particle's acceleration.

Example 3.2. The continuous parameter t can take all real values. Write down the derivatives $\mathrm{d}\mathbf{r}/\mathrm{d}t$ and $\mathrm{d}^2\mathbf{r}/\mathrm{d}t^2$ for the vector $\mathbf{r} = (\sin t)\mathbf{i} + t\mathbf{j}$. Also, sketch the curve whose parametric equation is $\mathbf{r} = \mathbf{r}(t)$.

The first and second derivatives are

$$\begin{aligned} \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} &= (\cos t)\mathbf{i} + \mathbf{j}, \\ \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} &= (-\sin t)\mathbf{i}. \end{aligned}$$

The sketch is shown in Fig. 3.1.

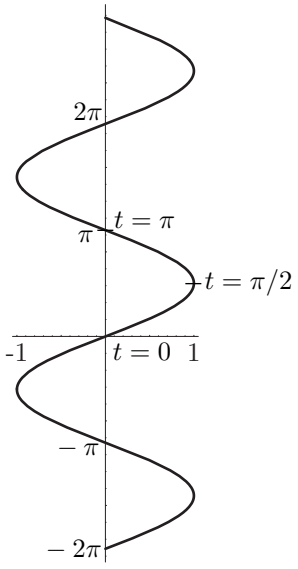

Figure 3.1: Sketch of the curve defined parametrically by $\mathbf{r} = (\sin t)\mathbf{i} + t\mathbf{j}$

It is easy to prove, by writing out the components and collecting terms, that if $\mathbf{F}$ and $\mathbf{G}$ are vector functions of u, then

$$\frac{\mathrm{d}(\mathbf{F}.\mathbf{G})}{\mathrm{d}u} = \mathbf{F}.\frac{\mathrm{d}\mathbf{G}}{\mathrm{d}u} + \frac{\mathrm{d}\mathbf{F}}{\mathrm{d}u}.\mathbf{G}.$$

Proof:

$$\begin{aligned} \frac{\mathrm{d}(\mathbf{F}.\mathbf{G})}{\mathrm{d}u} &= \frac{\mathrm{d}}{\mathrm{d}u}(f_1g_1 + f_2g_2 + f_3g_3) \\ &= f_1\frac{\mathrm{d}g_1}{\mathrm{d}u} + f_2\frac{\mathrm{d}g_2}{\mathrm{d}u} + f_3\frac{\mathrm{d}g_3}{\mathrm{d}u} + \frac{\mathrm{d}f_1}{\mathrm{d}u}g_1 + \frac{\mathrm{d}f_2}{\mathrm{d}u}g_2 + \frac{\mathrm{d}f_3}{\mathrm{d}u}g_3 \\ &= \mathbf{F}.\frac{\mathrm{d}\mathbf{G}}{\mathrm{d}u} + \frac{\mathrm{d}\mathbf{F}}{\mathrm{d}u}.\mathbf{G}. \quad \text{Q.E.D.} \end{aligned}$$

It's also straightforward to show that cross products work the same way. **Exercise 3.1.** Sketch the curves whose parametric equations are

(a) $\mathbf{r} = (3\sin\pi t)\mathbf{i} + (2\cos\pi t)\mathbf{j}$

(b) $\mathbf{r} = (\cos\pi t)\mathbf{j}$

(c) $\mathbf{r} = t\mathbf{i} + t^2\mathbf{k}$

($-\infty \le t \le \infty$), and write down the derivatives $d\mathbf{r}/dt$ and $d^2\mathbf{r}/dt^2$ where they are defined. □

If $\mathbf{F}$ is a vector function of more than one variable, say $\mathbf{F} = \mathbf{F}(u, v)$, then it is straightforward to define its partial derivatives with respect to u or v, in terms of partial derivatives of its components. Thus, for example, if $\mathbf{F} = (f_1(u, v), f_2(u, v), f_3(u, v))$, then

$$\frac{\partial \mathbf{F}}{\partial u} = \left(\frac{\partial f_1}{\partial u}, \frac{\partial f_2}{\partial u}, \frac{\partial f_3}{\partial u}\right).$$

We have already met an example of this for the surface $\mathbf{r} = \mathbf{r}(u, v)$ in Chapter 2.

3.2 Vector Fields

(See Thomas 16.2)
For the rest of this course, we shall be concerned mostly with vectors and scalars which depend on position in three-dimensional space, i.e. which are functions of three variables x, y, z. We have already met a function $f(x,y,z)$ where f is one number (a scalar); from here on, this will be called a **scalar field**, where the word "field" means that it is a function of (x,y,z), and the "scalar" means the function value at each point is a scalar.

(**Note** : Sometimes things may depend also on a fourth variable, such as time t, or we may only be interested in their values on a particular path $\mathbf{r}(s)$ where s is a parameter; but this doesn't change the key results.)

A vector depending on position in 3-D space is said to constitute a **vector field**. We write a vector $\mathbf{F}$ that varies with position as

$$\mathbf{F} = \mathbf{F}(x,y,z) \equiv \mathbf{F}(\mathbf{r})$$

An example is shown in Figure 3.2. In order to actually specify a vector field $\mathbf{F}$, we need to write it out in terms of its components, each depending on position, so

$$\begin{aligned} \mathbf{F} &= (F_1(x,y,z),\ F_2(x,y,z),\ F_3(x,y,z)) \\ &= F_1(x,y,z)\,\mathbf{i} + F_2(x,y,z)\,\mathbf{j} + F_3(x,y,z)\,\mathbf{k} \quad . \end{aligned}$$

Clearly this is rather cumbersome so we'll often write $\mathbf{F}(\mathbf{r})$ or just $\mathbf{F}$; but remember to actually calculate things you'll often need to write it out in full.

We have already met one example of a vector field: given a scalar field U, we have defined the gradient as

$$\nabla U = \frac{\partial U}{\partial x}\mathbf{i} + \frac{\partial U}{\partial y}\mathbf{j} + \frac{\partial U}{\partial z}\mathbf{k} \quad .$$

Here ∇U is itself a vector, and it (usually) depends on position, so it is actually a vector field.

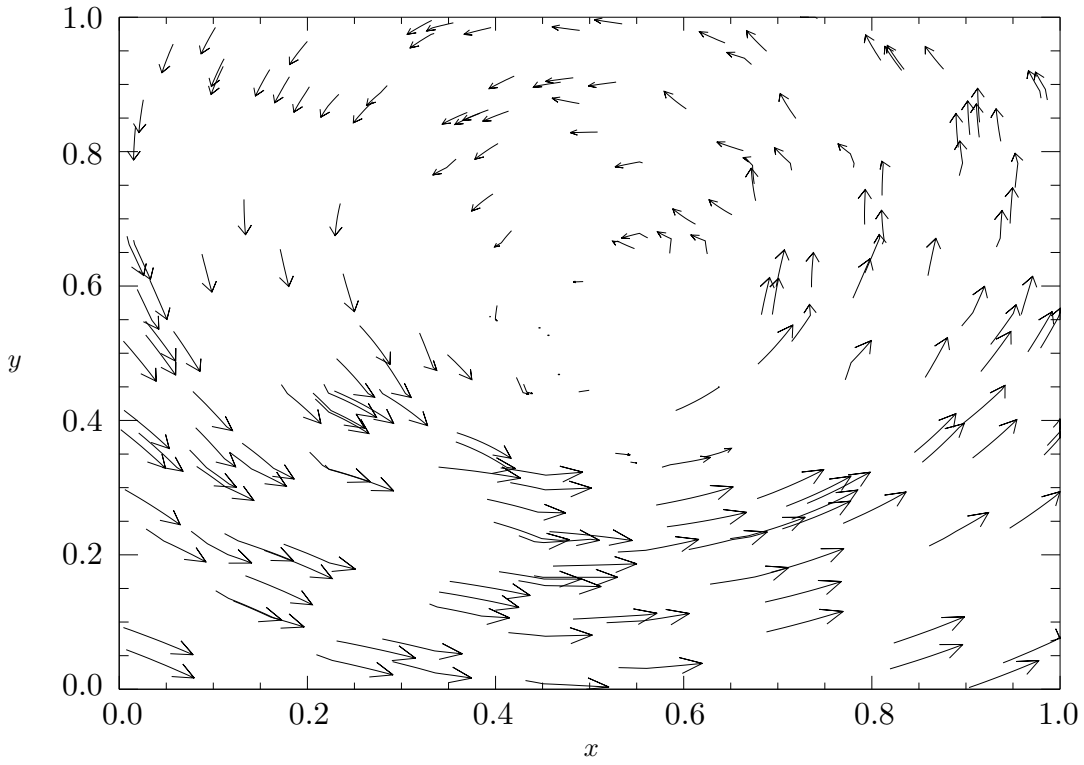

Figure 3.2: Example of a flow. In this case the speed and direction at each point is a function of the position (x,y)

A physical example of a vector field is the velocity in a flowing fluid (e.g. the water in the oceans, moving because of currents and tides; or the air in the atmosphere, moving because of winds). The velocity at any point in the fluid is a vector quantity – it has magnitude and direction. If we attach a velocity vector to each point of the flowing fluid, we have a vector field defined in the region occupied by the fluid.

Another physical example is a magnetic field; now things are not necessarily moving with time, but the magnetic field has a direction and a strength at each point in space; so at each point in space we have a vector; and this vector (in general) varies with position so it is a vector field.

We can add vector fields and multiply them by a constant in the obvious way, so if $\mathbf{F}$ and $\mathbf{G}$ are two vector fields then $\mathbf{F}+\mathbf{G}$ is also a vector field, and if λ is a constant then $\lambda\mathbf{F}$ is also a vector field.

Given a vector field, we could of course now differentiate the vector field with respect to each of the coordinates (x,y,z) in turn, in the manner described in the previous section; this gives us a total of 9 partial derivatives

$$\frac{\partial F_1}{\partial x},\ \frac{\partial F_2}{\partial x},\ldots,\frac{\partial F_1}{\partial y},\ldots,\frac{\partial F_3}{\partial z}$$

(In this course, we will be assuming that $\mathbf{F}$ is a smoothly-varying function of position, so all these derivatives exist at all points of interest $\mathbf{r}$, except possibly for one or two singular points).

Note: the set of all 9 derivatives of a component by a coordinate forms a quantity of a new kind, called a **tensor**. These are used in fluid dynamics, solid mechanics and relativity, for example. However, in this course we will **not** deal with tensors, we will restrict ourselves to forming scalar and vector quantities from these 9 derivatives. To do this, it will turn out that we have to take certain special combinations which are "well behaved" if we rotate the x,y,z axes; these will turn out to be forming the dot and cross products of ∇

with **F** , where

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$$

is the operator called "**del**" which we met previously in forming the gradient of a scalar. Note again that ∇ is not a true vector (because on its own we can't define its length or direction), but it is a vector differential operator.

3.3 The Divergence of a vector field

(See Thomas 16.8)
Suppose $\mathbf{F}(x,y,z) = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ is a vector field. The divergence of **F**, written div **F** or $\nabla \cdot \mathbf{F}$ is defined to be

$$\nabla \cdot \mathbf{F} \equiv \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \tag{3.1}$$

Here div **F** is a scalar (there are no $\mathbf{i}, \mathbf{j}, \mathbf{k}$'s in the result) and generally depends on position, so it is a scalar field.

We can also get the above result if we write out ∇ and **F** in components,

$$\nabla \cdot \mathbf{F} = (\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}) \cdot (F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k})$$

and write out all 9 terms then use the properties $\mathbf{i}.\mathbf{i} = 1$, $\mathbf{i}.\mathbf{j} = 0$, etc.

Note that, given a scalar field f, we found a vector field ∇f. Here, given a vector field **F**, we have produced a scalar field $\nabla \cdot \mathbf{F}$.

We can also write $\nabla \cdot \mathbf{F}$ as div **F**. These notations are completely interchangeable.

It is easy to show, by direct calculation, that div behaves as expected for addition and multiplication by a constant λ, i.e.

$$\nabla \cdot (\mathbf{F} + \mathbf{G}) = (\nabla \cdot \mathbf{F}) + (\nabla \cdot \mathbf{G}) \qquad ,$$
$$\nabla \cdot (\lambda \mathbf{F}) = \lambda (\nabla \cdot \mathbf{F})$$

The geometrical meaning of the divergence is as follows: consider a point **r** and consider a small closed surface surrounding that point: if the divergence div **F** is positive at **r**, then on average the vector field **F** is pointing "away" from the point and out of the surface. If the divergence is negative, then on balance **F** is pointing towards the point and into the surface. (See Fig. 3.3.) This idea will be made precise when we come to the Divergence Theorem in the next Chapter.

A vector field **F** for which $\nabla \cdot \mathbf{F} = 0$ everywhere is called **divergence-free** or **solenoidal**. The reason for the name **solenoidal** is historical: that a solenoid is a coiled wire that produces a magnetic field, and a magnetic field **B** is an example of a field that has $\nabla \cdot \mathbf{B} = 0$ everywhere (this is an observational fact, and arises because magnetic monopoles have never been found in many searches).

Example 3.3. If $\mathbf{F} = 3xy^2\mathbf{i} + e^z\mathbf{j} + xy\sin z\mathbf{k}$, calculate $\nabla \cdot \mathbf{F}$.

$$\nabla \cdot \mathbf{F} = \frac{\partial(3xy^2)}{\partial x} + \frac{\partial e^z}{\partial y} + \frac{\partial(xy\sin z)}{\partial z} = 3y^2 + xy\cos z.$$

Exercise 3.2. If $\mathbf{F} = (y-x)\mathbf{i} + (z-y)\mathbf{j} + (x-z)\mathbf{k}$, calculate $\nabla \cdot \mathbf{F}$. [Answer: -3] □

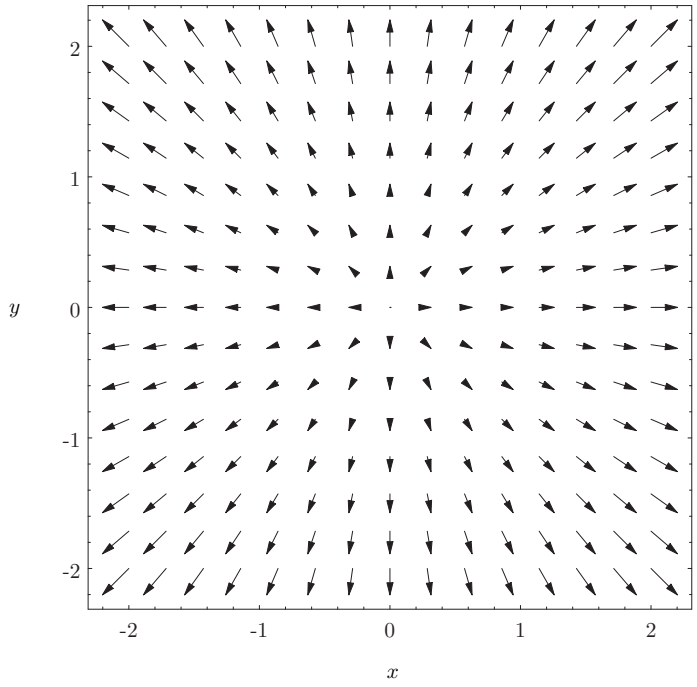

Figure 3.3: Example of a vector field with positive divergence (everywhere): $\mathbf{F} = x\mathbf{i} + y\mathbf{j}$.

3.4 The Curl of a vector field

(See Thomas 16.7)
The curl of a vector field **F** is defined to be

$$\nabla \times \mathbf{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k}. \tag{3.2}$$

Note that curl **F** is a vector, since there are $\mathbf{i}, \mathbf{j}, \mathbf{k}$ on the RHS; and it generally depends on position so it's a new **vector** field.

We can write $\nabla \times \mathbf{F}$ as curl **F** – again the two notations are completely interchangeable. It is convenient to remember $\nabla \times \mathbf{F}$ in terms of a determinant like the one for $\mathbf{v} \times \mathbf{w}$:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ F_1 & F_2 & F_3 \end{vmatrix}.$$

It is easy to verify, by writing out the determinant in full, that this is equivalent to the original definition.

It is also easy to show, by writing out the components, that if $\mathbf{F}, \mathbf{G}$ are any two vector fields,

$$\nabla \times (\mathbf{F} + \mathbf{G}) = (\nabla \times \mathbf{F}) + (\nabla \times \mathbf{G})$$

and if λ is any constant then

$$\nabla \times (\lambda \mathbf{F}) = \lambda (\nabla \times \mathbf{F})$$

Note that the equality above **only** works if λ is a constant (independent of x, y, z): see the next section for more general products.

The geometrical meaning of the curl is as follows. Loosely speaking, if at some point in space the component of the curl in the **n** direction is positive, it means that in the vicinity of the point and in a plane

normal to **n**, the vector field tends to go round in an anticlockwise direction if one looks along vector **n**. If the component of the curl were negative, it would mean that the vector field tends to go round in a clockwise direction. (See Fig. 3.4.) This idea will be made more precise when we come to Stokes's Theorem.

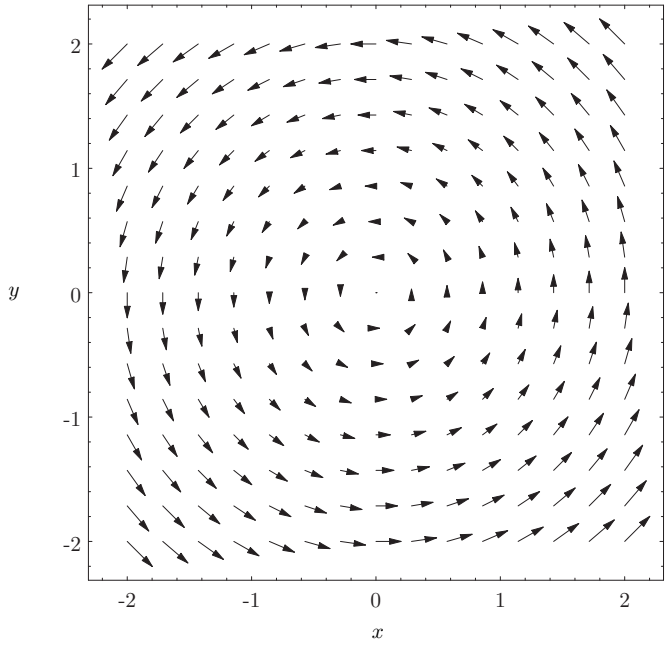

Figure 3.4: Example of a vector field with positive curl (in the z direction): $\mathbf{F} = x\mathbf{j} - y\mathbf{i}$.

A vector field **F** for which $\nabla \times \mathbf{F} = \mathbf{0}$ everywhere is called *curl-free* or *irrotational*.

Example 3.4. The velocity in a fluid is $\mathbf{v} = y\mathbf{i} - x\mathbf{j} + 0\mathbf{k}$. Find $\nabla \times \mathbf{v}$.

$$\nabla \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ y & -x & 0 \end{vmatrix} = \mathbf{i}(0-0) + \mathbf{j}(0-0) + \mathbf{k}(-1-1) = -2\mathbf{k}.$$

Exercise 3.3. If $\mathbf{F} = (x^2+y^2+z^2)\mathbf{i} + (x^4 - y^2z^2)\mathbf{j} + xyz\mathbf{k}$, find $\nabla \times \mathbf{F}$. □

Exercise 3.4. Find the divergence $(\nabla \cdot \mathbf{F})$ and curl $(\nabla \times \mathbf{F})$ of the following vector fields:

$\mathbf{F} = x^2\mathbf{i} + xz\mathbf{j} - 3z\mathbf{k}$

$\mathbf{F} = x^2\mathbf{i} - 2xy\mathbf{j} + 3xz\mathbf{k}$

$\mathbf{F} = \nabla(1/r)$ where $r = (x^2+y^2+z^2)^{1/2} \neq 0$.

□

3.5 Grad, Div and Curl of products

(See Thomas 16.7 and the exercises to 16.8)
We can now consider the application of grad, div and curl to products. We saw above that grad, div and curl behave in the "obvious" way for addition and multiplication by a constant.

However, we can also multiply scalar and/or vector fields together (in pairs) to get new scalar and vector fields. Altogether there are four ways to do this, as follows: if we have two scalar fields $U(\mathbf{r}), V(\mathbf{r})$ we can ordinary-multiply them (at each point $\mathbf{r}$) to get a new scalar field $UV = U(\mathbf{r})V(\mathbf{r})$; likewise for a scalar field $U(\mathbf{r})$ and a vector field $\mathbf{F}(\mathbf{r})$ we can use ordinary multiplication to give $U\mathbf{F} \equiv U(\mathbf{r})\mathbf{F}(\mathbf{r})$; the value is a vector, so this is a vector field. Also, if we have two vector fields $\mathbf{F}, \mathbf{G}$ we can define their dot product $\mathbf{F}.\mathbf{G} \equiv (\mathbf{F}(\mathbf{r}) \cdot (\mathbf{G}(\mathbf{r}))$ and their cross product $\mathbf{F} \times \mathbf{G}$ in the obvious way, by taking the dot or cross products of each field at the **same** point $\mathbf{r}$. Clearly $\mathbf{F} \cdot \mathbf{G}$ is a scalar field, and $\mathbf{F} \times \mathbf{G}$ is a vector field.

Note: in each of these products, the values of $U, V, \mathbf{F}, \mathbf{G}$ are taken at the **same** point $\mathbf{r}$ in the product. In longer equations, it is common to not bother writing in all the $(\mathbf{r})$'s, because if something is defined as a field then we know it is a function of $\mathbf{r}$.

We can now apply grad, div and curl to these products, but only for the following allowed combinations: to apply grad, we have to have a product which is itself a scalar field: that can be either an ordinary multiple of two scalar fields, say UV, or a scalar product (dot product) of two vector fields, $\mathbf{F} \cdot \mathbf{G}$.

Div and curl can only be applied to a vector field, so the possible products we could have look like $U\mathbf{F}$ or the cross product $\mathbf{F} \times \mathbf{G}$ above.

If we were dealing with functions of a single variable, the derivative would just give the well-known product rule for derivatives,

$$\frac{\mathrm{d}(fg)}{\mathrm{d}x} = f\frac{\mathrm{d}g}{\mathrm{d}x} + \frac{\mathrm{d}f}{\mathrm{d}x}g. \tag{3.3}$$

Some of the vector cases are just like that, but some are more complicated: we next give the results, and discuss the details afterwards. There are six cases as we've outlined above (two each for grad, div and curl).

For grad of products we have:

$$\begin{aligned} \nabla(UV) &= U(\nabla V) + V(\nabla U) && (3.4)\\ \text{or} \quad \mathrm{grad}\,(UV) &= U\,\mathrm{grad}\,V + V\,\mathrm{grad}\,U \\ \nabla(\mathbf{F}\cdot\mathbf{G}) &= \mathbf{F}\times(\nabla\times\mathbf{G}) + \mathbf{G}\times(\nabla\times\mathbf{F}) + (\mathbf{F}.\nabla)\mathbf{G} + (\mathbf{G}\cdot\nabla)\mathbf{F} && (3.5) \end{aligned}$$

For div of products we have:

$$\begin{aligned} \nabla\cdot(U\mathbf{F}) &= U(\nabla\cdot\mathbf{F}) + (\nabla U).\mathbf{F} && (3.6)\\ \nabla\cdot(\mathbf{F}\times\mathbf{G}) &= \mathbf{G}\cdot(\nabla\times\mathbf{F}) - \mathbf{F}\cdot(\nabla\times\mathbf{G}) && (3.7) \end{aligned}$$

and for curl of products, we have:

$$\begin{aligned} \nabla\times(U\mathbf{F}) &= U(\nabla\times\mathbf{F}) + (\nabla U)\times\mathbf{F} && (3.8)\\ &= U(\nabla\times\mathbf{F}) - \mathbf{F}\times(\nabla U) \\ \nabla\times(\mathbf{F}\times\mathbf{G}) &= \mathbf{F}(\nabla\cdot\mathbf{G}) + (\mathbf{G}.\nabla)\mathbf{F} - \mathbf{G}(\nabla\cdot\mathbf{F}) - (\mathbf{F}.\nabla)\mathbf{G} && (3.9) \end{aligned}$$

We see above that equations 3.4, 3.6, 3.7 and 3.8 look quite similar to 3.3, except for the minus sign in 3.7 and the possible minus sign in 3.8.

Note also that Eqs. 3.4 and 3.5 are symmetrical in the two variables, while 3.7 and 3.9 are antisymmetric, i.e. they must change sign if $\mathbf{F}, \mathbf{G}$ are swapped, due to the antisymmetry of the cross product.

Note: if you set $U(x,y,z) = \lambda$ =constant in the above, that is a (very boring) but legal scalar field with $\nabla U = \mathbf{0}$ everywhere; then you'll see Eqs. 3.4, 3.6, 3.8 reduce to the obvious cases of multiplication by a constant which we've met before. But for multiplication by a non-constant scalar U, the second terms involving ∇U appear on the RHS.

The other two equations 3.5, 3.9 are more complicated, and involve the new operator $(\mathbf{G}.\nabla)$: this is defined so for a scalar field V, if $\mathbf{G} = (G_1, G_2, G_3)$,

$$(\mathbf{G}.\nabla)V = \left(G_1\frac{\partial}{\partial x} + G_2\frac{\partial}{\partial y} + G_3\frac{\partial}{\partial z}\right)V = G_1\frac{\partial V}{\partial x} + G_2\frac{\partial V}{\partial y} + G_3\frac{\partial V}{\partial z}\ ,$$

For a vector field $\mathbf{F}$, the notation $(\mathbf{G}.\nabla)\mathbf{F}$ is to be interpreted as $(\mathbf{G}.\nabla F_1, \mathbf{G}.\nabla F_2, \mathbf{G}.\nabla F_3)$, taking $\mathbf{F} = (F_1, F_2, F_3)$; Thus writing out the whole thing, we have

$$(\mathbf{G}.\nabla)\mathbf{F} = \left(G_1\frac{\partial F_1}{\partial x} + G_2\frac{\partial F_1}{\partial y} + G_3\frac{\partial F_1}{\partial z},\ G_1\frac{\partial F_2}{\partial x} + G_2\frac{\partial F_2}{\partial y} + G_3\frac{\partial F_2}{\partial z},\ G_1\frac{\partial F_3}{\partial x} + G_2\frac{\partial F_3}{\partial y} + G_3\frac{\partial F_3}{\partial z}\right)$$

This is essentially the directional derivative of vector $\mathbf{F}$ in the direction of $\mathbf{G}$, i.e. it is $|\mathbf{G}|$ times the derivative $d\mathbf{F}/ds$ along the direction of the unit vector parallel to $\mathbf{G}$.

(Warning: the form of this definition will not persist in curvilinear coordinates, but the directional derivative will remain the same).

Note: you are not expected to memorise Eqs. 3.5 and 3.9, but you may be given those formulae in an exam question. You should know the definition of $(\mathbf{G}.\nabla)\mathbf{F}$ above.

Example 3.5. Let $\mathbf{a}$ be a constant vector, and $r = |\mathbf{r}|$ as usual. Then, using Eq 3.8,

$$\begin{aligned}\nabla \times (r\mathbf{a}) &= r(\nabla \times \mathbf{a}) - \mathbf{a} \times \nabla r \\ &= \mathbf{0} - \frac{\mathbf{a} \times \mathbf{r}}{r}\end{aligned}$$

since the curl of a constant $\mathbf{a}$ is zero, and $\nabla r = \mathbf{r}/r$ (as in Coursework 2).

Example 3.6. Let $\mathbf{a}$ be a constant vector. Then, using Equation 3.9,

$$\begin{aligned}\nabla \times (\mathbf{a} \times \mathbf{r}) &= \mathbf{a}(\nabla \cdot \mathbf{r}) + (\mathbf{r}.\nabla)\mathbf{a} - \mathbf{r}(\nabla \cdot \mathbf{a}) - (\mathbf{a}.\nabla)\mathbf{r} \\ &= 3\mathbf{a} + \mathbf{0} - \mathbf{0} - \mathbf{a} \\ &= 2\mathbf{a}\end{aligned}$$

(On the top line, the two middle terms differentiate the constant $\mathbf{a}$ so are both zero, and it is simple to check from the definitions that $\nabla \cdot \mathbf{r} = 3$ and $(\mathbf{a}.\nabla)\mathbf{r} = \mathbf{a}$.)

Proofs:

All of the equations 3.4 to 3.9 can be proved directly from the definitions by inserting components, expanding out using the ordinary derivative-of-product rule and doing some rearrangement; this can be fairly long, but is not difficult.

For a couple of examples: firstly for Eq. 3.4 it is simple, we have

$$\nabla(UV) = \mathbf{i}\frac{\partial}{\partial x}(UV) + \mathbf{j}\frac{\partial}{\partial y}(UV) + \mathbf{k}\frac{\partial}{\partial z}(UV)$$

$$
\begin{aligned}
&= \mathbf{i}(U\frac{\partial V}{\partial x}+V\frac{\partial U}{\partial x})+\mathbf{j}(U\frac{\partial V}{\partial y}+V\frac{\partial U}{\partial y})+\mathbf{k}(U\frac{\partial V}{\partial z}+V\frac{\partial U}{\partial z}) \\
&= U\left(\mathbf{i}\frac{\partial V}{\partial x}+\mathbf{j}\frac{\partial V}{\partial y}+\mathbf{k}\frac{\partial V}{\partial z}\right)+V\left(\mathbf{i}\frac{\partial U}{\partial x}+\mathbf{j}\frac{\partial U}{\partial y}+\mathbf{k}\frac{\partial U}{\partial z}\right) \\
&= U(\nabla V)+V(\nabla U) \qquad \text{QED.}
\end{aligned}
$$

Next we'll prove Eq.3.8: the product $U\mathbf{F}$ is a vector field with components (UF_1, UF_2, UF_3); inserting those into the definition of curl,

$$
\begin{aligned}
\nabla\times(U\mathbf{F}) &= \mathbf{i}\left(\frac{\partial}{\partial y}(UF_3)-\frac{\partial}{\partial z}(UF_2)\right)+\mathbf{j}\left(\frac{\partial}{\partial z}(UF_1)-\frac{\partial}{\partial x}(UF_3)\right)+\mathbf{k}\left(\frac{\partial}{\partial x}(UF_2)-\frac{\partial}{\partial y}(UF_1)\right) \\
&= \mathbf{i}\left(U\frac{\partial F_3}{\partial y}+F_3\frac{\partial U}{\partial y}-U\frac{\partial F_2}{\partial z}-F_2\frac{\partial U}{\partial z}\right)+\mathbf{j}\left(U\frac{\partial F_1}{\partial z}+F_1\frac{\partial U}{\partial z}-U\frac{\partial F_3}{\partial x}-F_3\frac{\partial U}{\partial x}\right) \\
&\quad +\mathbf{k}\left(U\frac{\partial F_2}{\partial x}+F_2\frac{\partial U}{\partial x}-U\frac{\partial F_1}{\partial y}-F_1\frac{\partial U}{\partial y}\right)
\end{aligned}
$$

Now we just re-order the 12 terms so that the six with a $U\partial F_i$ come first, then the six with an $F_i\partial U$ come next; and from the definitions, it becomes clear that the result is

$$
\nabla\times(U\mathbf{F}) = U(\nabla\times\mathbf{F})+(\nabla U)\times\mathbf{F} \qquad \text{QED.}
$$

The others can be proved in a similar way, though it gets quite long for Eqs. 3.5 and 3.9. Much shorter proofs can be given using **index notation**, but this is no longer on the syllabus.

Note: As always, be careful what is a scalar and what is a vector. Remember that in an equation

$$
(\textit{expression } 1) = (\textit{expression } 2)+(\textit{expression } 3)+\ldots
$$

expressions 1,2,3 ... must be either all scalars or all vectors, since you cannot add a scalar and a vector. Check that you understand that in the above equations, all the expressions are vectors for 3.4, 3.5, 3.8, 3.9 (because grad() or curl() give a vector result); while they are scalars for 3.6 and 3.7 because div gives a scalar result.

3.6 Vector second derivatives: applying ∇ twice

We also have a second set of identities arising from applying **two** of grad, div or curl in succession. Here grad U and curl **F** produce vector fields, to which either div or curl can be applied; while div **F** produces a scalar field, and then we can apply grad to that. This gives a total of five allowed cases, which are as follows:

$$
\begin{aligned}
\text{div}(\text{grad U}) &= \nabla\cdot(\nabla U)\equiv\nabla^2 U && (3.10)\\
\text{curl}(\text{grad } U) &= \nabla\times(\nabla U)=\mathbf{0} && (3.11)\\
\text{div}(\text{curl } \mathbf{F}) &= \nabla\cdot(\nabla\times\mathbf{F})=0 && (3.12)\\
\text{curl}(\text{curl } \mathbf{F}) &= \nabla\times(\nabla\times\mathbf{F})=\nabla(\nabla\cdot\mathbf{F})-\nabla^2\mathbf{F} && (3.13)\\
\text{grad}(\text{div } \mathbf{F}) &= \nabla(\nabla\cdot\mathbf{F})=\nabla\times(\nabla\times\mathbf{F})+\nabla^2\mathbf{F} && (3.14)
\end{aligned}
$$

We see here that two of these cases (curl grad U, and div curl **F**) are identically zero ; this is true for any fields, as long as they are sufficiently well behaved that the partial derivatives commute, see below. These

two zero cases can be helpfully memorised by the fact that they would also give zero if ∇ was replaced by an ordinary vector **a** ; but beware, this sort of rule is not applicable to every equation containing ∇.

The first equation above Eq. 3.10 introduces a new operator ∇^2 called the **Laplacian**; this is very important in a wide range of physical problems, and we will meet it extensively in Chapter 7. In components, combining the definition of grad U from Chapter 1 and plugging that into Eq. 3.1, we get simply

$$\nabla^2 U \equiv \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} \tag{3.15}$$

This Laplacian operator can be applied to either a scalar field or a vector field, producing a field of the same type; in the above, U is a scalar field and $\nabla^2 U$ is another scalar field.

For a vector field **F**, to get $\nabla^2\mathbf{F}$ we apply ∇^2 to each component of **F** separately, giving

$$\nabla^2\mathbf{F} = \mathbf{i}\nabla^2 F_1 + \mathbf{j}\nabla^2 F_2 + \mathbf{k}\nabla^2 F_3$$

so $\nabla^2\mathbf{F}$ is another vector field.

Also note that the last two of the above equations 3.13 and 3.14 are just a rearrangement of each other, giving a relationship between curl curl **F**, grad div **F** and $\nabla^2\mathbf{F}$.

All of the relations above can be proved by direct substitution, e.g.:

Proof of 3.11:

$$\begin{aligned} \text{curl}(\nabla U) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial U/\partial x & \partial U/\partial y & \partial U/\partial z \end{vmatrix} \\ &= \left(\frac{\partial^2 U}{\partial y \partial z} - \frac{\partial^2 U}{\partial z \partial y}, \ \frac{\partial^2 U}{\partial z \partial x} - \frac{\partial^2 U}{\partial x \partial z}, \ \frac{\partial^2 U}{\partial x \partial y} - \frac{\partial^2 U}{\partial y \partial x} \right) = 0. \end{aligned}$$

[Note, we assume that the function U is sufficiently well-behaved for its partial second derivatives to commute.]

The relation curl grad $U = \mathbf{0}$ is particularly useful, since it is often interesting to ask, given some vector field **F**, can we find a scalar field U such that $\nabla U = \mathbf{F}$? If we can, this simplifies things from 3 functions of position to 1 function.

Now we can show that if our given **F** has curl $\mathbf{F} \neq \mathbf{0}$, it is not possible to find such a scalar field U, as follows: choose any scalar field U, and define a vector field $\mathbf{H} = \nabla U$. We'd like to find a U such that $\mathbf{H} = \mathbf{F}$. But from Eq. 3.11, $\nabla \times \mathbf{H} = \nabla \times (\nabla U) = \mathbf{0}$. Therefore $\mathbf{H} \neq \mathbf{F}$: so, if curl $\mathbf{F} \neq \mathbf{0}$ then it is **not** possible to express **F** as the gradient of any scalar field U.

The converse is also true: we will show in the next chapter that if curl $\mathbf{F} = \mathbf{0}$ everywhere in a given domain, then we *can* find a scalar field U with $\nabla U = \mathbf{F}$: and we'll also show how to construct the desired U using a suitable integral. This requires vector integration, which we'll do in the next Chapter.

Chapter 4

Vector integrals and integral theorems

Last revised: 1 Nov 2010.

Syllabus covered:
1. Line, surface and volume integrals.
2. Vector and scalar forms of Divergence and Stokes's theorems. Conservative fields: equivalence to curl-free and existence of scalar potential. Green's theorem in the plane.

Calculus I and II covered integrals in one, two and three dimensional Euclidean (flat) space (i.e. $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}^3$). We are still working in $\mathbb{R}^3$ so there is no generalization to be applied to volume or triple integrals, but we will generalise one dimensional integration from a straight line to an integral along a curve, and we will generalise two-dimensional integration from a region in a plane to a curved surface.

We will also be working with integration of vectors, though in many cases we will be using a scalar product so the final quantity to be integrated becomes a scalar. In the cases with a scalar product:

$\int f(x)\,\mathrm{d}x$ generalizes to $\int_{\mathscr{C}} \mathbf{F}\cdot\mathrm{d}\mathbf{r}$ on a curve $\mathscr{C}$, called a **line integral** (section 4.1).

$\int\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ generalizes to $\int_{\mathscr{S}} \mathbf{F}\cdot\mathrm{d}\mathbf{S}$ over a surface $\mathscr{S}$, called a **surface integral** (section 4.2).

We will then have to study the generalizations of

$$\int_a^b \frac{\mathrm{d}f}{\mathrm{d}x}\,dx = f(b) - f(a)\,, \tag{4.1}$$

called the 'fundamental theorem of calculus', which we use in the proofs. This theorem relates a one-dimensional integral to a (pair of) zero-dimensional evaluations at the two endpoints $x = a, b$. The higher dimensional versions do the following:

Stokes's theorem relates the surface integral of a curl to a line integral (2 dimensions to 1) around the edge of the surface: see section 4.6.

The Divergence Theorem[1] relates the volume integral of a divergence to a surface integral (3 dimensions to 2) over the boundary of the volume: see section 4.4.

There is also a special case of Stokes's theorem where the surface is a plane: this is Green's theorem

[1] First discovered by Joseph Louis Lagrange in 1762, then independently rediscovered by Carl Friedrich Gauss in 1813, by George Green in 1825 and in 1831 by Mikhail Vasilievich Ostrogradsky, who also gave the first proof of the theorem. Thus the result may be called Gauss's Theorem, Green's theorem, or Ostrogradsky's theorem.

relating the integral of a curl to a line integral (2 dimensions to 1): see section 4.5.

[Aside: All these are in fact special cases of the general Stokes's theorem which relates an $n-1$ dimensional integral of a field to the n dimensional integral of its derivative. Here the field is a generalization of a vector field called an $(n-1)$-form field.]

Finally we will discuss the application to potentials, and the proofs.

Before moving on to line and surface integrals, we consider the case where one wants to integrate a vector function $\mathbf{F}(u)$ of one variable, u, with respect to u. The integral can be calculated simply by integrating the components (in Cartesian coordinates) of $\mathbf{F} = (F_1, F_2, F_3)$:

$$\int \mathbf{F}\,\mathrm{d}u = \left(\int F_1\,\mathrm{d}u, \int F_2\,\mathrm{d}u, \int F_3\,\mathrm{d}u\right) \tag{4.2}$$

$$= \mathbf{i}\int F_1\,\mathrm{d}u + \mathbf{j}\int F_2\,\mathrm{d}u + \mathbf{k}\int F_3\,\mathrm{d}u \tag{4.3}$$

Integration of a vector in this case is just a set of three ordinary integrals. The restriction to Cartesian coordinates can be overcome by looking at the definition in vectorial terms: we go back to the basic definition of integration, which leads to a geometrical picture of $\mathbf{G} \equiv \int_a^b \mathbf{F}\mathrm{d}u$ (see Fig. 4.1):

$$\mathbf{G} = \int_a^b \mathbf{F}\,\mathrm{d}u = \lim_{\delta u_p \to 0} \sum_{p=1}^{N} \mathbf{F}(u)\delta u_p \ .$$

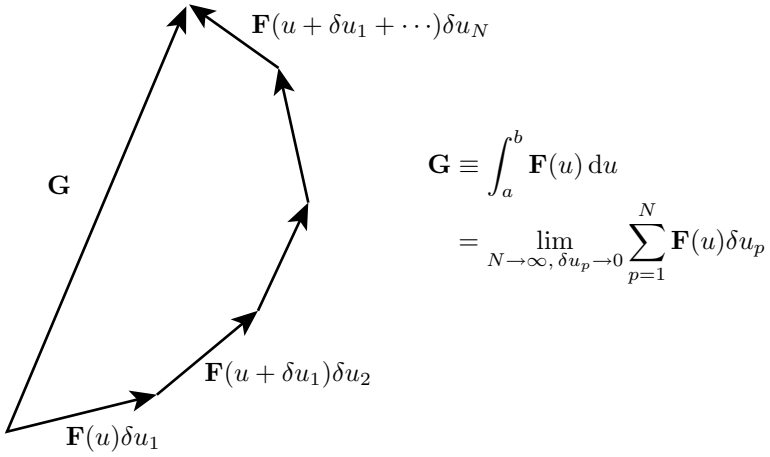

Figure 4.1: Geometrical picture of $\mathbf{G} = \int_a^b \mathbf{F}\,\mathrm{d}u = \lim_{\delta u_p\to 0}\sum_{p=1}^{N}\mathbf{F}(u)\delta u_p$.

Example 4.1. If $\mathbf{v}(t) \equiv \mathrm{d}\mathbf{r}/\mathrm{d}t$ is the velocity of a particle, as a function of time t, then

$$\int_{t_1}^{t_2} \mathbf{v}\,\mathrm{d}t = \int_{t_1}^{t_2} \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t = \int_{\mathbf{r}(t_1)}^{\mathbf{r}(t_2)} \mathrm{d}\mathbf{r} = \mathbf{r}(t_2) - \mathbf{r}(t_1) \qquad .$$

Note here that $\mathbf{v}$ is the vector velocity of the particle, so the time-integral is the vector distance between the two end-points. If we had put v instead of $\mathbf{v}$ in the integral, then the result would be a scalar equal to the total arc-length of the curved path $\mathbf{r}(t)$, as we met in Chapter 2.

Warning: there seems to be a common belief that an integral always represents an area or volume. This comes from 1-D integration where $\int f(x)\,dx$ can be shown as an area between a curve $y = f(x)$ and the $x-$axis; or in 2-D integration the result $\int h(x,y)\,dxdy$ can be expressed as a volume between the xy plane

and the surface $z = h(x,y)$. However, when we take general line and surface integrals the results are not necessarily areas and volumes; we will see that these integrals can represent various things such as distance travelled, work done by a force, flow of a fluid crossing a surface, etc, but there is not always a simple geometrical picture for the result of an integral.

4.1 Line Integrals

(See Thomas 16.1 and 16.2: note that Thomas begins by defining a scalar integral $\int f|\mathrm{d}\mathbf{r}|$, in the notation below. I come back to this at the end of this section.)
Suppose $\mathbf{F} = (F_1, F_2, F_3)$ is a vector field defined in some region of space, and $\mathscr{C}$ is a parametrized curve through that region from $\mathbf{r}_1$ to $\mathbf{r}_2$, so that $\mathscr{C}$ is given by

$$\mathbf{r}(t) = (g(t), h(t), q(t)) \qquad (t_1 \le t \le t_2),$$

and $\mathbf{r}_1 = \mathbf{r}(t_1)$, $\mathbf{r}_2 = \mathbf{r}(t_2)$. Then, one can define the line integral

$$\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

to be

$$\int_{t_1}^{t_2} \mathbf{F}(\mathbf{r}(t)) \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t \quad \equiv \quad \int_{t_1}^{t_2} \left(F_1 \frac{\mathrm{d}g}{\mathrm{d}t} + F_2 \frac{\mathrm{d}h}{\mathrm{d}t} + F_3 \frac{\mathrm{d}q}{\mathrm{d}t} \right) \mathrm{d}t \tag{4.4}$$

Warning: do not forget to write the components of $\mathbf{F}$ in terms of the parameter t, so that t is the only variable that appears inside the integral!. Hence you must write $\mathbf{F}(\mathbf{r}) = \mathbf{F}(\mathbf{r}(t))$, so we replace $F_1(x, y, z)$ by $F_1(g(t), h(t), q(t))$, and so on; then we evaluate the dot product of $\mathbf{F}$ and $d\mathbf{r}/dt$, before finally integrating over t to get the numerical answer.

Second warning: it seems to be easy to confuse where one has to use $\mathbf{r}(t)$ and where one uses $\mathrm{d}\mathbf{r}/\mathrm{d}t$; you have to evaluate $\mathbf{F}$ at position $\mathbf{r}(t)$, while the line-segment $d\mathbf{r}$ is given by $(d\mathbf{r}/dt)\,dt$.

The above is just a version of the fundamental definition of an integral as the limit of lots of small contributions. In this case it's the scalar products of $\mathbf{F}(\mathbf{r})$ with small displacements $d\mathbf{r}$ along $\mathscr{C}$:

$$\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \lim_{\delta\mathbf{r}_p \to 0} \sum_{p=1}^{N} \mathbf{F}(\mathbf{r}) \cdot \delta\mathbf{r}_p$$

If we are given a geometrical description of the curve without a parametrization, we have to first **find a parametrisation** of the described curve to actually evaluate the integral. For lines, circles, ellipses and so we can use, for example, (1.33) and (1.18)–(1.20).

Example 4.2. Evaluate the integral $\int \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ for the vector field $\mathbf{F} = -4xy\mathbf{i} + 8y\mathbf{j} + 2\mathbf{k}$, from the origin to the point $(2, 4, 1)$ along the following three paths:

1. along the curve $\mathbf{r} = t\mathbf{i} + t^2\mathbf{j} + \frac{1}{2}t\mathbf{k}$, $0 \le t \le 2$,
2. from the origin to $(2, 0, 0)$, then from there to $(2, 4, 0)$, then to $(2, 4, 1)$, along straight lines [Note that the answer will be the sum of the three parts: a path may have several pieces, providing the next one begins where the previous one ends.]
3. on the surface $4x^2 + y^2 = 32z$ along a line with constant y/x.

Note that only for the first of these do we have the parametrization given: in the second and third we'll have to make a parametrization from the definitions.

1. In this case we are given the parametrised curve $\mathbf{r}(t)$ as above, and from that we get $\dfrac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{i} + 2t\mathbf{j} + \frac{1}{2}\mathbf{k}$, and $\mathbf{F}(\mathbf{r}(t)) = -4(t)(t^2)\mathbf{i} + 8(t^2)\mathbf{j} + 2\mathbf{k}$. The final bits we need is the $t-$ values at the given endpoints $\mathbf{r}_1 = (0,0,0)$ and $\mathbf{r}_2 = (2,4,1)$; it is easy to see those are $t = 0$ and $t = 2$ (solve the easiest equation e.g. $x = t$, and plug that in to the other two to check). Putting things together, we have

$$\begin{aligned}
\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F}.\mathrm{d}\mathbf{r} &= \int_{t=0}^{2} \mathbf{F}(\mathbf{r}(t)).\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t \\
&= \int_0^2 (-4t^3\mathbf{i} + 8t^2\mathbf{j} + 2\mathbf{k}).(\mathbf{i} + 2t\mathbf{j} + \tfrac{1}{2}\mathbf{k})\,\mathrm{d}t \\
&= \int_0^2 \left[(-4t^3)(1) + (8t^2)(2t) + (2)(\tfrac{1}{2})\right]\,\mathrm{d}t \\
&= \int_0^2 (12t^3 + 1)\,\mathrm{d}t \\
&= \left[3t^4 + t\right]_0^2 = 48 + 2 = 50.
\end{aligned}$$

2. Now our given "curve" is three straight line segments joined end-to-end and we need parametrisations for each, separately.
   The first segment is from $(0,0,0)$ to $(2,0,0)$. The straight line is, from the general form of Eq. 1.33 for the case of a line joining $\mathbf{r}_1$ and $\mathbf{r}_2$, i.e. $\mathbf{r} = \mathbf{r}_1 + t(\mathbf{r}_2 - \mathbf{r}_1)$,

   $$\mathbf{r} = \mathbf{0} + t(2\mathbf{i}), \quad 0 \le t \le 1,$$

   Here we could call $2t$ simply x, so

   $$\mathbf{r} = x\mathbf{i}, \quad 0 \le x \le 2.$$

   Along this line we have $\mathrm{d}\mathbf{r} = \mathbf{i}\,\mathrm{d}x$. To get the value of $\mathbf{F}$ we substitute $y = z = 0$ into the general form for $\mathbf{F}$, giving $\mathbf{F} = 2\mathbf{k}$. Taking the scalar product, $\mathbf{F}\cdot\mathrm{d}\mathbf{r} = 0$ and hence this segment gives a zero integral.

   In the second segment, from $(2,0,0)$ to $(2,4,0)$, we similarly get

   $$\mathbf{r} = 2\mathbf{i} + y\mathbf{j}, \quad 0 \le y \le 4$$

   so along it, $\mathrm{d}\mathbf{r} = \mathbf{j}\mathrm{d}y$. Substituting $x = 2$, $z = 0$ in $\mathbf{F}$ we have $\mathbf{F} = -8y\mathbf{i} + 8y\mathbf{j} + 2\mathbf{k}$. So $\mathbf{F}\cdot\mathrm{d}\mathbf{r} = 8y\,\mathrm{d}y$ and hence this gives

   $$\int_0^4 8y\,\mathrm{d}y = [4y^2]_0^4 = 64$$

   In the last segment, from $(2,4,0)$ to $(2,4,1)$,

   $$\mathbf{r} = 2\mathbf{i} + 4\mathbf{j} + z\mathbf{k}, \quad 0 \le z \le 1$$

   so along it we have $\mathrm{d}\mathbf{r} = \mathbf{k}\,\mathrm{d}z$, and $x = 2$, $y = 4$ gives $\mathbf{F} = -32\mathbf{i} + 32\mathbf{j} + 2\mathbf{k}$, so $\mathbf{F}\cdot\mathrm{d}\mathbf{r} = 2\,\mathrm{d}z$ and hence this gives $\int_0^1 2\,\mathrm{d}z = 2$.

   Finally adding the integrals from the three segments together, we get the full line integral over our given path $= 0 + 64 + 2 = 66$.

3. Now we are integrating along a line in a curved surface; the equation for the line is not given explicitly, but we are told two things which let us solve for it: the line is in the surface $4x^2 + y^2 = 32z$, and our line has constant y/x so $y = kx$ for some constant k. Geometrically, our line will be the intersection of a plane $y = kx$ (containing the $z-$axis) with the above surface. Since at the second end point $x = 2$ and $y = 4$, we need $k = 2$ so $y = 2x$. Substituting that in $4x^2 + y^2 = 32z$ gives $8x^2 = 32z$ so $x = 2\sqrt{z}$, and

$y = 4\sqrt{z}$. Now we have both x and y in terms of z, so we can use z as the one parameter for our curve: we have

$$\mathbf{r} = 2\sqrt{z}\mathbf{i} + 4\sqrt{z}\mathbf{j} + z\mathbf{k} \qquad 0 \le z \le 1 \quad ,$$

where the limits on z follow from the given endpoints. Once we have a one-parameter expression for the curve $\mathbf{r}(z)$, it is straightforward: we get $\mathrm{d}\mathbf{r} = (\mathbf{i}/\sqrt{z} + 2\mathbf{j}/\sqrt{z} + \mathbf{k})\mathrm{d}z$, while $\mathbf{F}(\mathbf{r}(z)) = -32z\mathbf{i} + 32\sqrt{z}\mathbf{j} + 2\mathbf{k}$. Hence inserting those into Eq. 4.4, remembering to take the scalar product, we have

$$\int_0^1 (-32\sqrt{z} + 64 + 2)\,\mathrm{d}z = \int_0^1 (66 - 32\sqrt{z})\,\mathrm{d}z = [66z - 64z^{3/2}/3]_0^1 = 66 - \frac{64}{3} \quad ,$$

which we could also write as $44\frac{2}{3}$.

Note: in the above, we could alternatively have chosen x as the one parameter, and write $y = 2x$, $z = x^2/4$ to get $\mathbf{r} = x\mathbf{i} + 2x\mathbf{j} + (x^2/4)\mathbf{k}$, and range $0 \le x \le 2$. It is straightforward to check that this gives the same result $44\frac{2}{3}$ for the line integral.

As well as giving some examples of how to calculate line integrals, this example makes the important point that in general the result depends on the curve, not just on its two endpoints. We shall return to this matter in Section 4.7, where we will find that if $\mathbf{F}$ has zero curl (irrotational), the resulting line integral only depends on the two endpoints, not the curve between them.

Exercise 4.1. Calculate $\int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r}$, where $\mathbf{F} = 4yz\mathbf{i} - 3z\mathbf{j} + 2x^2\mathbf{k}$, over each of the following curves from $(0,0,0)$ to $(1,1,1)$:

(a) C: $\mathbf{r} = t\mathbf{i} + t\mathbf{j} + t\mathbf{k} \quad 0 \le t \le 1$

(b) C: $\mathbf{r} = t^2\mathbf{i} + t\mathbf{j} + t^3\mathbf{k} \quad 0 \le t \le 1$

□

If the vector field $\mathbf{F}$ represents a **force** (e.g. gravitational force), then

$$\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F}.\mathrm{d}\mathbf{r}$$

is called a **work integral** and its value is the **work done** by the force for a particle moving between $\mathbf{r}_1$ and $\mathbf{r}_2$, which equals the increase in energy of the body acted on. This occurs because for each small movement $d\mathbf{r}$, (small enough to be a straight line), if θ is the local angle between $\mathbf{F}$ and $d\mathbf{r}$, then $F\cos\theta$ is the component of force *parallel* to $d\mathbf{r}$, so $\mathbf{F}\cdot d\mathbf{r} = F\,dr\cos\theta$ is the work done by the force, along the small step $d\mathbf{r}$. The line integral just adds up that work along all the small steps along the path, so the line integral is the total work done from $\mathbf{r}_1$ to $\mathbf{r}_2$.

If instead of representing a force, $\mathbf{F}$ represents the velocity field in a fluid, and if $\mathscr{C}$ is some curve in the fluid, then $\int_{\mathscr{C}} \mathbf{F}\cdot\mathrm{d}\mathbf{r}$ is called the **flow** along curve $\mathscr{C}$. If $\mathscr{C}$ is a closed curve, the flow is called the **circulation** around $\mathscr{C}$.

Finally, note that Thomas's form $\int f|\mathrm{d}\mathbf{r}|$ is obtained if one assumes that $\mathbf{F}$ is parallel to the unit tangent vector to the curve, $\mathbf{t} = \dfrac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\Big/\left|\dfrac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right|$, at all points on the curve, since $\dfrac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\mathrm{d}t = \mathbf{t}|\mathrm{d}\mathbf{r}|$, and in this case, taking $\mathbf{F} = f\mathbf{t}$,

$$\mathbf{F}\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t = f\mathbf{t}\cdot\mathbf{t}|\mathrm{d}\mathbf{r}| = f|\mathrm{d}\mathbf{r}| \qquad .$$

Thus Thomas's starting point is simply a special case of the general line integral.

4.2 Surface integrals

(See Thomas 16.5 and 16.6, but be aware that Thomas starts by defining the integral of a scalar, using what is, in the notation below, $\int f|\mathrm{d}\mathbf{S}|$.)

To define surface integrals, we now have to take into account that a small area on a curved surface has both a magnitude and a direction (the normal to the surface) associated with it, so we can represent a small area as a vector, as we saw in Chapter 2.

Consider an area S in a plane (see Fig. 4.2a). If $\mathbf{n}$ is a unit vector perpendicular to the plane, then the vector representing the area, $\mathbf{S}$, is defined to be

$$\mathbf{S} = S\mathbf{n}$$

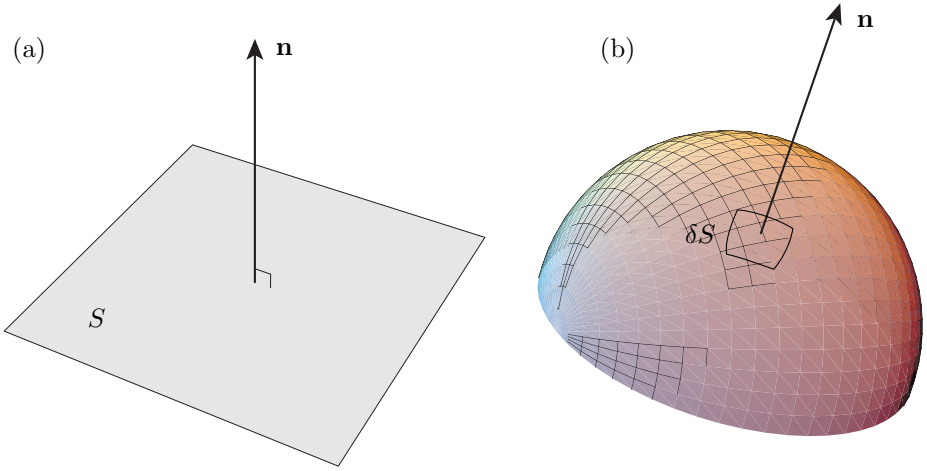

Figure 4.2: (a) Normal $\mathbf{n}$ to a plane area S. The vector area is $\mathbf{S} = S\mathbf{n}$. (b) Normal $\mathbf{n}$ to a more general surface. The vector area of the small surface element is $\delta\mathbf{S} = \delta S\,\mathbf{n}$, where δS is the magnitude of the area.

In the case of a curved surface in three dimensions (see 4.2b), we need to pick a small area δS which is small enough to be approximated as (almost) flat, and define the vector $\delta\mathbf{S}$ for that area element δS as

$$\delta\mathbf{S} = \delta S\,\mathbf{n} \qquad ,$$

where $\mathbf{n}$ is a unit vector normal to the surface element δS. Note we are still using the convention that vectors are written in bold type and the same symbol in ordinary type means the magnitude, thus $\delta S = |\delta\mathbf{S}|$. In the limit we shall write $\mathrm{d}S$ rather than δS. (Thomas uses $\mathrm{d}\sigma$ for this $\mathrm{d}S$.)

Note we still have a sign ambiguity in this definition, because either direction of the unit vector along the normal line could be used. One case where we can fix the sign is the case of a **closed** surface, where $\mathbf{n}$ is generally taken to be the **outward-pointing** unit normal vector. If the surface is not closed, we will have to explicitly specify geometrically one of the two possible directions for $\mathbf{n}$.

Now that we have defined how to represent a small area as a vector, we can now define the **surface integral** for a vector field $\mathbf{F}$ over a general curved surface $\mathscr{S}$:

$$\int_{\mathscr{S}} \mathbf{F}\cdot\mathrm{d}\mathbf{S} \quad = \quad \int_{\mathscr{S}} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S \qquad . \tag{4.5}$$

Such an integral is also called the **flux** of $\mathbf{F}$ across area $\mathscr{S}$. Since the quantity integrated is a scalar product of two vectors, the answer is a scalar quantity.

These surface integrals arise in a number of physical situations: one example is the case where **F** represents the velocity field in a fluid, where the surface integral represents the volume of fluid crossing the surface $\mathscr{S}$ per unit time. Another example is if **F** is a magnetic field **B**, in which case the integral would be the *magnetic flux* across the surface $\mathscr{S}$. (These results occur because $F\cos\theta$ is the component of **F** parallel to the local normal **n** i.e. perpendicular to the surface; while the component of **F** parallel to the surface (perpendicular to **n**) does not contribute to the flux across the surface. Thus, the flux of **F** crossing any small patch of surface $d\mathbf{S}$ is $|\mathbf{F}|\cos\theta\, dS$, which is $\mathbf{F}\cdot d\mathbf{S}$ from the definition of the dot product. The integration then just adds up the contribution from all the infinitesimal patches, to get the flux crossing the whole curved surface.

The double integrals in a plane that we met before, $\int\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y$, can be thought of as integrals of $\mathbf{F}.\mathrm{d}\mathbf{S}$, where
$\mathbf{F} = f\mathbf{k}$ and $\mathrm{d}\mathbf{S} = (\mathrm{d}x\,\mathrm{d}y)\mathbf{k}$.

The tricky part is, once we are given a field **F** and a surface $\mathscr{S}$, to turn the general form $\int_{\mathscr{S}}\mathbf{F}.\mathrm{d}\mathbf{S}$ into a double integral that we can actually do. We shall give some general rules after studying some examples.

We next look at three examples of increasing difficulty: one is a simple plane case, the second a curved surface where the integral is easy, and the third gives us the patterns we need for the general case.

Example 4.3. If $\mathbf{F} = (3x,\, 2xz,\, 3)$, evaluate the flux of **F** across the surface $\mathscr{S}$: $z = 0$, $0 \le x \le 1$, $0 \le y \le 2$ (where the normal is to be in the positive z direction).

Here the given surface is a rectangle in the xy-plane, so the normal **n** is $\pm\mathbf{k}$. We are told to take the plus sign. We need to integrate over x and y with limits as above:

$$\int_{\mathscr{S}}\mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S = \int_0^1\int_0^2 (3x\mathbf{i}+(2x)0\mathbf{j}+3\mathbf{k}).(0\mathbf{i}+0\mathbf{j}+1\mathbf{k})\,\mathrm{d}y\,\mathrm{d}x = \int_{x=0}^1\int_{y=0}^2 3\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 6\,\mathrm{d}x = 6.$$

Example 4.4. If the velocity field of a fluid is $\mathbf{v} = \frac{1}{r^2}\mathbf{e}_r$, where r is the distance from the origin O and $\mathbf{e}_r$ is a unit vector at position **r** pointing away from the origin, find the flux $\int\mathbf{v}\cdot\mathbf{n}\,\mathrm{d}S$ across a sphere $\mathscr{S}$ of radius a whose centre is at the origin. (The outward normal should be taken.)

In this case, the outward normal and $\mathbf{e}_r$ are the same vector, so

$$\mathbf{v}\cdot\mathbf{n} = \frac{1}{r^2}\mathbf{e}_r\cdot\mathbf{e}_r = \frac{1}{r^2}$$

($\mathbf{e}_r.\mathbf{e}_r = 1$ because $\mathbf{e}_r$ is a unit vector). On the given sphere of radius a, $r = a$, so

$$\int_{\mathscr{S}}\mathbf{v}.\mathbf{n}\,\mathrm{d}S = \int_{\mathscr{S}}\frac{1}{a^2}\,\mathrm{d}S = \frac{1}{a^2}\times(\text{Area of sphere of radius } a) = \frac{1}{a^2}4\pi a^2 = 4\pi \quad .$$

using the fact that $\frac{1}{a^2}$ is a constant, so can be taken outside the integral sign.

Example 4.5. Find the flux of the field $\mathbf{F} = z\mathbf{k}$ across the portion of the sphere $x^2+y^2+z^2 = a^2$ in the first octant (this is the $1/8$-th of space in which x, y and z are all ≥ 0) with normal taken in the direction away from the origin.

This example is easier in spherical polars (see later), but we can do it in Cartesians. Write the required part of the sphere as a surface $z = \sqrt{a^2-x^2-y^2}$ (note that for a whole sphere we would also need the points where $z = -\sqrt{a^2-x^2-y^2}$, the square root being understood to be the non-negative one). Consider the displacement vector for a small change $\mathrm{d}x$, by taking the derivative of $\mathbf{r} = (x,\, y,\, \sqrt{a^2-x^2-y^2})$ as in section

3.1. It will be

$$\frac{\partial \mathbf{r}}{\partial x}\mathrm{d}x = \left(\frac{\partial x}{\partial x}, \frac{\partial y}{\partial x}, \frac{\partial z}{\partial x}\right)\mathrm{d}x = \left(1, 0, \frac{-x}{\sqrt{a^2-x^2-y^2}}\right)\mathrm{d}x \tag{4.6}$$

and similarly a small change in y gives a displacement

$$\frac{\partial \mathbf{r}}{\partial y}\mathrm{d}y = \left(0, 1, \frac{-y}{\sqrt{a^2-x^2-y^2}}\right)\mathrm{d}y\,. \tag{4.7}$$

The magnitude of the corresponding area element is then given by the area of a parallellogram with sides (4.6) and (4.7), and the normal direction is perpendicular to them both, so we need their cross-product

$$\begin{aligned}\mathrm{d}\mathbf{S} &= \left(1, 0, \frac{-x}{\sqrt{a^2-x^2-y^2}}\right)\mathrm{d}x \times \left(0, 1, \frac{-y}{\sqrt{a^2-x^2-y^2}}\right)\mathrm{d}y \\ &= \left(\frac{x}{\sqrt{a^2-x^2-y^2}}\mathbf{i} + \frac{y}{\sqrt{a^2-x^2-y^2}}\mathbf{j} + \mathbf{k}\right)\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

Thus $\mathbf{F}\cdot\mathrm{d}\mathbf{S} = z\,\mathrm{d}x\,\mathrm{d}y = \sqrt{a^2-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y$.

Now we need the limits on the variables. The first octant of the sphere lies above the first quadrant of the circle $x^2+y^2=a^2$, $z=0$, so we will have

$$\int_{x=0}^{a}\int_{y=0}^{\sqrt{a^2-x^2}} \sqrt{a^2-x^2-y^2}\,\mathrm{d}y\,\mathrm{d}x\,.$$

The rest of the problem is just a double integral like those in Calculus II. We can do it by a substitution such as $y=\sqrt{a^2-x^2}\sin\xi$ which gives

$$\int_0^a (a^2-x^2)\int_{\xi=0}^{\pi/2}\cos^2\xi\;\mathrm{d}\xi\,\mathrm{d}x$$

and this turns out to be $\pi a^3/6$ using the double-angle formula.

Note that parametrization by a pair of coordinates will not always give all the surface: for example, consider the surface consisting of two touching perpendicular squares, one square with a vertex at the origin and sides 1 along the x and y axes, and the similar square in the (x,z) plane: this surface cannot be covered by any pair of the Cartesian coordinates, though it can easily be split into two pieces each of which separately can be handled that way, and the results added.

The final part of the above example provides general methods for turning a surface integral like Eq. 4.5 into a double integral we can actually do. We next look at 3 cases:

1. Surface given by two parameters $\mathbf{r}(u,v)$.
2. Surface given by $z=h(x,y)$
3. Surface given by $g(x,y,z)=$const.

Note that if we are only given a geometrical "description" of the surface, we will need to put our surface into one of the above forms before we proceed: which is easiest may depend on the surface, but usually the two-parameter case is simplest.

4.2.1 Surface integral: surface given by two parameters

First consider the case where *the surface is given, or can be found, in terms of two parameters*: several examples were covered in Chapter 2. See also Thomas 16.6, and diagrams 16.55 and 16.56. For a surface given by two parameters u, v we have:

$$\mathbf{r}(u,v) = x(u,v)\mathbf{i} + y(u,v)\mathbf{j} + z(u,v)\mathbf{k} \ .$$

Now we can do the surface integral as follows:

1. Calculate the partial derivatives $\dfrac{\partial \mathbf{r}}{\partial u}$ and $\dfrac{\partial \mathbf{r}}{\partial v}$.
2. Calculate the cross product
$$\mathrm{d}\mathbf{S} = \left(\frac{\partial \mathbf{r}}{\partial u}\right) \times \left(\frac{\partial \mathbf{r}}{\partial v}\right) du\, dv$$
As we showed previously, this vector is normal to the surface and has magnitude equal to the area of the small parallelogram with four corners given by $\mathbf{r}(u,v)$, $\mathbf{r}(u+du,v)$, $\mathbf{r}(u,v+dv)$, $\mathbf{r}(u+du,v+dv)$, so it is the $d\mathbf{S}$ we want.
3. Express $\mathbf{F}$ in terms of u, v using $\mathbf{r} = \mathbf{r}(u,v)$ as given above and substituting.
4. Form the scalar product $\mathbf{F} \cdot \mathrm{d}\mathbf{S}$
5. From the given geometry of the surface, work out appropriate limits on u, v and perform the double integral over du and dv.

This gives us finally

$$\int_{\mathscr{S}} \mathbf{F} \cdot d\mathbf{S} = \int_v \int_u \mathbf{F}(\mathbf{r}(u,v)) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) du\, dv$$

Warning: note that the cross-product above may be opposite to the required normal direction, so one may need to take its negative (which is equivalent to just swapping the order in the cross-product).
Both for this reason, and for working out limits on the variables, it is a good idea to draw a sketch first.

For the standard surfaces such as cylinders, spheres and ellipsoids we already know some parametrizations, (1.28)–(1.31).

4.2.2 Surface integral: surface $z = h(x,y)$

The second case to consider is where we have a surface given as one coordinate is a function of the other two, e.g. $z = h(x,y)$. This is essentially a special case of the more general two-parameter case above where $x = u,\ y = v,\ z = h(u,v)$. Just using x and y as the parameters, we get the surface as $\mathbf{r} = (x, y, h(x,y))$, and partial differentiation gives

$$\frac{\partial \mathbf{r}}{\partial x} = (1, 0, \partial h/\partial x) \qquad \frac{\partial \mathbf{r}}{\partial y} = (0, 1, \partial h/\partial y) \qquad ;$$

so the area element on the curved surface $z = h(x,y)$ is again the cross product of the above, which is

$$\mathrm{d}\mathbf{S} = (-\partial h/\partial x, -\partial h/\partial y, 1)\, \mathrm{d}x\mathrm{d}y \qquad .$$

Next we evaluate $\mathbf{F}(\mathbf{r})$ on the surface using $\mathbf{r} = (x, y, h(x,y))$ again, we evaluate the scalar product $\mathbf{F} \cdot d\mathbf{S}$, and finally do the double integral with respect to x, y.

(There are other similar cases if instead x is given as a function of y,z by $x = g(y,z)$; this is very similar to the above except for swapping x,y,z) .

Aside: It is also useful to note that the unit normal to the surface $z = h(x,y)$ is

$$\mathbf{n} = \frac{1}{\sqrt{(\partial h/\partial x)^2 + (\partial h/\partial y)^2 + 1}}(-\partial h/\partial x, -\partial h/\partial y, 1)$$

Since $\mathbf{n}$ is a unit vector, the angle θ this makes with the z axis is given by

$$\cos\theta = \mathbf{k}.\mathbf{n} = 1/\sqrt{(\partial h/\partial x)^2 + (\partial h/\partial y)^2 + 1} \qquad .$$

The magnitude $\mathrm{d}S = |\mathrm{d}\mathbf{S}|$ is then
$dS = \sqrt{(\partial h/\partial x)^2 + (\partial h/\partial y)^2 + 1}\,\mathrm{d}x\,\mathrm{d}y = \mathrm{d}x\,\mathrm{d}y/\cos\theta$.

This is not needed for the surface integral in the current case, but we will make use of this result in the next section .

4.2.3 Surface given by $g(x,y,z) = constant$

The third case of a surface integral is that where we are given a vector field $\mathbf{F}$, and where our surface is defined by a function $g(x,y,z) = \text{const}$, (and some specified boundaries), when we do not necessarily have a convenient parametrization. As long as the surface is single-valued in two coordinates, e.g. for a given x,y there is a unique z on the surface, we can use those two coordinates e.g. x,y as the two parameters as follows:

1. Calculate ∇g (which is the vector normal to the surface).
2. Find the unit normal vector in that direction $\mathbf{n} = \nabla g/|\nabla g|$.
3. Calculate $\cos\theta = \mathbf{n}\cdot\mathbf{k}$, where θ is the angle between $\mathbf{n}$ and the $+z$-direction.
4. Write $\mathrm{d}\mathbf{S} = \mathbf{n}\,\mathrm{d}S = \mathbf{n}\,\mathrm{d}x\,\mathrm{d}y/\cos\theta$, using the result from the previous subsection. (For a geometrical illustration, consider a 'light bulb' at $z = +\infty$. A small patch on our surface with area dS would cast a 'shadow' of area $dS\cos\theta$ on the xy plane; reversing this, the required area dS on the surface which casts a shadow of area $dx\,dy$ will be $dS = dx\,dy/\cos\theta$).

   Combining the above expressions for $\mathbf{n}$ and $\cos\theta$ gets us $\mathrm{d}\mathbf{S} = (\nabla g)\mathrm{d}x\,\mathrm{d}y/(\nabla g\cdot\mathbf{k})$.
5. Finally, use this to form $\mathbf{F}.\mathrm{d}\mathbf{S}$, and do the double integration with respect to x and y.

Thus, we can use (x, y) as our two parameters, provided $\cos\theta \neq 0$ over our range of x,y, and also provided that we can express $\mathbf{F}(\mathbf{r})$ on the surface in terms of x and y. Here we may need to solve for z in terms of x,y on our given surface; or if we are lucky, things may simplify so that at given x, y and $g(x,y,z)$ we can evaluate $\mathbf{F}.d\mathbf{S}$ without actually needing to solve for z.

Note that Thomas gives an even more general version of this where he considers a plane with normal $\mathbf{p}$ and an area $\mathrm{d}A$ in the plane (in place of $\mathbf{k}$ and $\mathrm{d}x\,\mathrm{d}y$): because he is working with $|\mathrm{d}S|$ he uses $|\cos\theta|$ and writes $1/|\cos\theta|$ as $|\nabla g|/|\nabla g.\mathbf{p}|$. While one is unlikely to need to use a general $\mathbf{p}$, that version has the advantage of covering the three cases $\mathbf{p} = \mathbf{i}$, $\mathbf{p} = \mathbf{j}$ and $\mathbf{p} = \mathbf{k}$ in one formula.

Exercise 4.2. If $\mathbf{F} = x\mathbf{i} + y\mathbf{j}$, evaluate

$$\int_S \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S$$

where S is the rectangular box formed by the six planes

$$x = 0, a\,, \quad y = 0, b\,, \quad z = 0, c\,.$$

□

Exercise 4.3. If $\mathbf{F} = 3y^2\mathbf{i} - \mathbf{j} + xz\mathbf{k}$, evaluate the integral $\int_{\mathscr{S}} \mathbf{F}.\mathrm{d}\mathbf{S}$, where $\mathscr{S}$ is the surface $z = 1$, $0 \le x \le 1$, $0 \le y \le x$ (take the normal pointing in the positive z direction).
[Answer: $1/3$] □

Exercise 4.4. If $\mathbf{F} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, evaluate

$$\int_S \mathbf{F} \cdot \mathbf{n}\, \mathrm{d}S$$

over the hemispherical surface S given by $z \ge 0$, $x^2 + y^2 + z^2 = a^2$, taking the normal outward from the origin.
[Answer: πa^2] □

To link up with Thomas, his initial $\int f|\mathrm{d}\mathbf{S}|$ is just $\int \mathbf{F}.\mathrm{d}\mathbf{S}$ for a vector field such that $\mathbf{F} = f\mathbf{n}$ on the surface.

4.3 Volume Integrals

In Cartesian coordinates, consider a small cuboid with one corner at (x,y,z) and sides (dx,dy,dz). This has the eight corners (x,y,z), $(x+dx,y,z)$, …, $(x+dx,y+dy,z+dz)$, and the infinitesimal volume of the cuboid is obviously $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$. Since in this course we will not be considering curved three-dimensional objects in four-dimensional space, we do not have to think about a vectorial version of dV.

However, the fact that dV is a *volume* element is an important way to look at it. If we re-label our space using new coordinates (u,v,w) , then taking small displacements du, dv, dw gives us small displacements $(\partial\mathbf{r}/\partial u)du$, $(\partial\mathbf{r}/\partial v)dv$, $(\partial\mathbf{r}/\partial w)dw$ in ordinary x,y,z space. These three vectors will form a small parallelepiped, and the volume of that parallelepiped dV is given by a scalar triple product of the three vectors above (see section 1.7); that will give the Jacobian determinant for change of variables in a triple integral,

$$dV = \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| du\,dv\,dw$$

as in section 1.3; so this explains why the Jacobian formula works.

Usually the integrand of a volume integral is a scalar. However, we could integrate vectors in $\mathbb{R}^3$, though this is not so often used. Given a vector field $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$, one can define

$$\int_V \mathbf{F}\,\mathrm{d}V = \left(\int_V F_1 \mathrm{d}V\right)\mathbf{i} + \left(\int_V F_2 \mathrm{d}V\right)\mathbf{j} + \left(\int_V F_3 \mathrm{d}V\right)\mathbf{k}$$

For example, $\mathbf{F}$ might be the momentum vector field in a fluid, (in that case we would have $\mathbf{F} = \rho\mathbf{v}$ where ρ is the density and $\mathbf{v}$ is the velocity); the volume integral above would then equal the total net momentum of that volume of fluid.

The most useful integrals we will deal with from here onwards are the line integral $\int_{\mathscr{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$, the flux across a surface, $\int_{\mathscr{S}} \mathbf{F} \cdot \mathrm{d}\mathbf{S}$, and the integral of a scalar over a volume, $\int_V f\,\mathrm{d}V$.

4.4 The Divergence Theorem

(See Thomas 16.8)
The Divergence Theorem states (following Thomas's wording) that "under suitable conditions":

Theorem 4.1 *The flux of a vector field* **F** *across a closed oriented surface* $\mathscr{S}$ *in the direction of the surface's outward unit normal vector field* **n** *equals the integral of* $\nabla\cdot\mathbf{F}$ *over the region* $\mathscr{D}$ *enclosed by the surface*

$$\int_{\mathscr{D}} \nabla\cdot\mathbf{F}\,\mathrm{d}V = \int_{\mathscr{S}} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S \equiv \int_{\mathscr{S}} \mathbf{F}\cdot\mathrm{d}\mathbf{S} \tag{4.8}$$

If asked to state this theorem, you **must** define the terms used, and state the conditions on the surface (i.e. "closed, oriented") and on the direction of the normal (outward).

We have not spelt out here in detail the 'suitable conditions' required of **F** and the surface. These, and a proof, are discussed in section 4.9, but will not be examinable.

Here the word 'Oriented' means we assign an outward direction for the normal to S in a consistent and continuous way. An S for which this is possible is called *orientable*: the Möbius strip (see Thomas Fig. 16.46) is an example of a non-orientable surface.

Note that it is not required that $\mathscr{S}$ has a single connected piece. For instance, it could have two parts, one inside the other, and then $\mathscr{D}$ would be the volume in between.

The Divergence theorem appears in a number of important physical situations such as Maxwell's equations in electromagnetism, and various cases in fluid dynamics. From a purely mathematical viewpoint, another use is that to calculate either of the integrals in it, we can use the other one if it is easier to do.

In the next example we calculate both sides of the Divergence Theorem for a simple case, and verify they really are equal.

Example 4.6. Suppose $f = xy$. Find a vector field **F** such that $\nabla\cdot\mathbf{F} = f$. Suppose V is the closed rectangular volume bounded by the planes $x = 0, a$, $y = 0, b$, $z = 0, c$, and $\mathscr{S}$ is the surface of the volume. Evaluate directly

$$\int_V f\,\mathrm{d}V \quad \text{and} \quad \int_{\mathscr{S}} \mathbf{F}.\mathbf{n}\,\mathrm{d}S$$

(where **n** is an outward normal), and show that they are equal – as they should be, according to the Divergence Theorem.

The volume integral is straightforward.

$$\begin{aligned}\int_0^c\int_0^b\int_0^a xy\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \int_0^c\int_0^b [\tfrac{1}{2}x^2y]_0^a\,\mathrm{d}y\,\mathrm{d}z = \int_0^c\int_0^b \tfrac{1}{2}a^2y\,\mathrm{d}y\,\mathrm{d}z \\ &= \int_0^c [\tfrac{1}{4}a^2y^2]_0^b\,\mathrm{d}z = \int_0^c \tfrac{1}{4}a^2b^2\,\mathrm{d}z = \tfrac{1}{4}a^2b^2[z]_0^c = \tfrac{1}{4}a^2b^2c.\end{aligned}$$

There are numerous ways to construct a vector field **F** of the required form, e.g. by integrating f with respect to x and making this the x-component of a vector **F**, so

$$\mathbf{F} = \left(x^2y/2\,,0,0\right)\ .$$

Our closed surface $\mathscr{S}$ enclosing V is a cuboid with six faces, so we must evaluate $\mathbf{F}.\mathbf{n}$ on each of the six and add the results. Since our cuboid is aligned with the x,y,z axes, on two of the faces, $\mathbf{n} = \pm\mathbf{i}$, on two $\mathbf{n} = \pm\mathbf{j}$ and on the last two $\mathbf{n} = \pm\mathbf{k}$.

Because $\mathbf{F} \propto \mathbf{i}$ is always parallel to the x-direction, $\mathbf{F}.\mathbf{n} = 0$ on the four faces where $\mathbf{n} = \pm\mathbf{j}, \pm\mathbf{k}$, so those give zero surface integral. The remaining faces are the two where $x = 0$ and $x = a$:
On the $x = 0$ face, $\mathbf{F} = \mathbf{0}$ and so $\mathbf{F}.\mathbf{n} = 0$. This leaves only the face $x = a$. On that face $\mathbf{F}.\mathbf{n} = (a^2y/2)\mathbf{i}.\mathbf{i} =$

$a^2y/2$, and we have $dS = dydz$. Integrating this over that face with respect to y,z gives

$$\int_{\mathscr{S}} \mathbf{F}.\mathbf{n} = \int_0^b \int_0^c \frac{1}{2}a^2 y \, \mathrm{d}z \, \mathrm{d}y = \left(\frac{1}{2}a^2\right)\left(\frac{1}{2}b^2\right)c = \frac{1}{4}a^2b^2c \ ,$$

which agrees with the volume integral of $\nabla \cdot \mathbf{F}$ above.

Example 4.7. A more typical example of the use of the Divergence Theorem is the following. Find the integral $\int_S \mathbf{A}.\mathrm{d}\mathbf{S}$ for $\mathbf{A} = (x, z, 0)$ and the surface S of a sphere of radius a.

Using the divergence theorem, the surface integral is equal to the volume integral $\int_V \nabla \cdot \mathbf{A}\, dV$ over the volume V interior to the sphere. But $\nabla \cdot \mathbf{A} = 1$, so the volume integral is $\int 1 \, \mathrm{d}V$ over the sphere, which is the volume of the sphere $= 4\pi a^3/3$.

Doing the surface integral $\int_S \mathbf{A} \cdot d\mathbf{S}$ directly is possible, but much more long-winded.

Example 4.8. Another good example is that from Example 4.5, where we evaluated a rather fiddly surface integral over 1/8th of a sphere. In that case, we were given $\mathbf{F} = z\mathbf{k}$; so $\nabla \cdot \mathbf{F} = 1$; and the Divergence theorem tells us that a volume integral of $\nabla \cdot \mathbf{F}$ is equal to the surface integral of $\mathbf{F} \cdot d\mathbf{S}$ over the *whole* surface bounding the volume. We may choose our volume as the interior of the 1/8 sphere, bounded by three planes $x = 0$, $y = 0$, $z = 0$ and the 1/8 sphere $x^2 + y^2 + z^2 = a^2$ with $x,y,z > 0$, then the volume integral of $\nabla \cdot \mathbf{F}$ is just $(1/8)$ (Volume of full sphere) $= \pi a^3/6$.

The surface integral is the sum of four parts: one part over the 1/8 surface of the sphere which we did before, plus three surface integrals over flat quarter-circles in each of the *xy*, *xz* and *yz* planes: those have *outward* unit normal vectors $-\mathbf{k}$, $-\mathbf{j}$, $-\mathbf{i}$ respectively since our volume is on the positive side of each plane. But $\mathbf{F} = z\mathbf{k}$, so for the second and third of those planes the dot product $\mathbf{F}.d\mathbf{S}$ is zero; and for the first plane, we are at $z = 0$ so $\mathbf{F} = 0$. Therefore, all three of the flat quarter-circles give us surface integrals of 0 ; so the surface integral of $\mathbf{F}.d\mathbf{S}$ over the 1/8 sphere is equal to the volume integral of $\nabla \cdot \mathbf{F}$, $= \pi a^3/6$, QED.

Exercise 4.5. State the Divergence Theorem. Evaluate both sides of the Divergence Theorem for the vector field $\mathbf{F} = xy^2z\mathbf{k}$ over a volume V which is the interior of the unit cube, i.e. the cube whose vertices are at $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$ and $(1, 1, 1)$. □

The Divergence Theorem equates two scalar values. However, one can derive from it vector identities. For example, we can obtain what is called the vector form of the theorem:

$$\int U \, \mathrm{d}\mathbf{S} = \int \nabla U \, \mathrm{d}V \qquad , \tag{4.9}$$

where U is a scalar field, and both sides of the above equation are vectors.

This is proved as follows: given the scalar field U, we choose any constant vector $\mathbf{a}$ and define a new vector field $\mathbf{F} = U\,\mathbf{a}$; next we apply the usual divergence theorem to $\mathbf{F}$, and the product rule Eq. 3.6 gives us $\nabla \cdot (U\mathbf{a}) = 0 + \mathbf{a} \cdot (\nabla U)$, so

$$\int \mathbf{a}U \cdot \, \mathrm{d}\mathbf{S} = \int \mathbf{a} \cdot (\nabla U) \, \mathrm{d}V \qquad .$$

Since $\mathbf{a}$ is a constant vector we can take it outside the integral signs, and finally choosing the cases $\mathbf{a}$ = $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ in turn, we prove Eq. 4.9.

4.5 Green's Theorem (in the plane)

(See Thomas 16.4: we take the statement he gives as Theorem 4, reworded. Note that the right side is a component of a curl.)

Theorem 4.2 *(Green's Theorem:) If $\mathscr{C}$ is a simple closed curve in the x-y plane, traversed counterclockwise, and M and N are suitably differentiable functions of x and y, then*

$$\int_{\mathscr{C}} (M\mathrm{d}x + N\mathrm{d}y) = \int\int_{\mathscr{R}} \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y,$$

where the area integral is over the region $\mathscr{R}$ enclosed by the curve $\mathscr{C}$.

Note that if asked to state the theorem you must state the nature of $\mathscr{C}$ ("simple closed") and the direction in which it is travelled.

Proof: The proof is an application of the Divergence Theorem, choosing a volume of height 1 in the z-direction above $\mathscr{R}$. (Or, if one proves Stokes's theorem first, of that theorem.) Take $\mathbf{F} = (N, -M, 0)$: then

$$\begin{aligned} \int (\nabla \cdot \mathbf{F}) \mathrm{d}V &= \int_{z=0}^{1} \int\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ &= \int\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y \end{aligned}$$

on integrating over z from 0 to 1. On the top and bottom of the volume, $\mathrm{d}\mathbf{S}$ is in the $\pm\mathbf{k}$ direction so $\mathbf{F}.\mathrm{d}\mathbf{S} = 0$. On the rest of the surface we have

$$\int \mathbf{F}.\mathrm{d}\mathbf{S} = \int\int N\mathrm{d}S_x - M\mathrm{d}S_y$$

where $\mathrm{d}S_x$ is the component of $\mathrm{d}\mathbf{S}$ along the x-axis. Using $\mathrm{d}\mathbf{r}_{\mathscr{C}} = (\mathrm{d}x, \mathrm{d}y, 0)$ along $\mathscr{C}$ and $\mathrm{d}\mathbf{r}_z = (0, 0, \mathrm{d}z)$ in the z-direction, $\mathrm{d}\mathbf{S} = \mathrm{d}\mathbf{r}_{\mathscr{C}} \times \mathrm{d}\mathbf{r}_z$ gives $\mathrm{d}S_x = \mathrm{d}y\,\mathrm{d}z$ and $\mathrm{d}S_y = -\mathrm{d}x\,\mathrm{d}z$, so

$$\begin{aligned} \int \mathbf{F}.\mathrm{d}\mathbf{S} &= \int\int_S N\ \mathrm{d}y\,\mathrm{d}z + \int\int_S M\ \mathrm{d}x\,\mathrm{d}z \\ &= \int_C N\ \mathrm{d}y + \int_C M\ \mathrm{d}x. \end{aligned}$$

where the second line follows because the z integral runs from 0 to 1 and the integrand is independent of z; now we have proved the two sides of the theorem are equal.

(Thomas's Theorem 3 is the same with N replaced by M and M replaced by $-N$. This version makes the right side look like a two-dimensional divergence. Sometimes you may see these called Green's theorem (first form) and Green's Theorem (second form) etc.)

Example 4.9. Use Green's theorem to evaluate

$$\int \left(xy\ \mathrm{d}y - y^2\ \mathrm{d}x\right)$$

around the unit square: straight path segments from the origin to $(1,0)$ to $(1,1)$ to $(0,1)$ and back to the origin.

In this case, $M = -y^2$ and $N = xy$; hence

$$\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = y + 2y = 3y$$

Thus the required integral is

$$\int_0^1 \int_0^1 3y \, \mathrm{d}y \, \mathrm{d}x = \int_0^1 (3/2) \, \mathrm{d}x = 3/2.$$

4.5.1 Area within a curve

From Green's Theorem, we can get a surprising expression for the area A inside a closed curve C bounding a region S in a plane is

$$\mathbf{n}A = \tfrac{1}{2} \oint_C \mathbf{r} \times \mathrm{d}\mathbf{r} \, ,$$

where $\mathbf{n}$ is the unit normal to the plane. We can assume without loss of generality that the plane of the curve is the x, y plane. Then $\mathbf{n} = \mathbf{k}$ and $\mathbf{r} \times \mathrm{d}\mathbf{r} \propto \mathbf{k}$, so we only need the z component of the integral which is

$$\tfrac{1}{2} \oint_C (x\,dy - y\,dx).$$

By Green's theorem in the plane this equals

$$\frac{1}{2} \int_S 2 \, dx\,dy = \int_S 1 \, dx\,dy = \text{ Area inside } C.$$

This can be useful for example if we are given a curve in parametric form $(x,y) = (f(t), g(t))$ which contains a closed loop, and we want the area of the loop: since the curve has a closed loop, then there are two values of t_1, t_2 where the curve returns to the same point, and (as long as the curve does not cross itself between t_1, t_2), we can evaluate the enclosed area within that loop using the above formula as

$$A \quad = \quad \frac{1}{2} \int_{t_1}^{t_2} x(t) \frac{dy}{dt} - y(t) \frac{dx}{dt} \, dt \tag{4.10}$$

A neat example of this is the case of the ellipse, $x = a\cos t, \quad y = b\sin t$; this clearly is a closed loop for $t_1 = 0, \; t_2 = 2\pi$, and we obtain the area as

$$A = \frac{1}{2} \int_0^{2\pi} (ab\cos^2 t + ab\sin^2 t) \, dt = \pi ab \qquad .$$

4.6 Stokes's Theorem

(See Thomas 16.7)
The other major theorem of similar character to the Divergence Theorem is Stokes's theorem which follows. (Because both are versions of the n-dimensional Stokes's theorem, we can prove Stokes's theorem from Green's and thence from the Divergence Theorem, which we do in section 4.9. It can also be proved directly.) We reword Thomas's version.

Theorem 4.3 *[Stokes's theorem]: If* $\mathbf{F}$ *is a (suitably differentiable) vector field, and* $\mathscr{C}$ *is a closed path bounding an oriented surface* $\mathscr{S}$*, then*

$$\int_{\mathscr{C}} \mathbf{F}.\mathrm{d}\mathbf{r} = \int_{\mathscr{S}} (\nabla \times \mathbf{F}).\mathbf{n} \, \mathrm{d}S \equiv \int_{\mathscr{S}} (\nabla \times \mathbf{F}).\mathrm{d}\mathbf{S}, \tag{4.11}$$

where $\mathscr{C}$ *is travelled counterclockwise with respect to the unit normal* $\mathbf{n}$ *of* $\mathscr{S}$ (i.e. counterclockwise as seen from the positive $\mathbf{n}$ side of $\mathscr{S}$).

Again note that if asked to state the theorem, you must state that $\mathscr{C}$ is closed, that it bounds $\mathscr{S}$, and that the directions of d$\mathbf{S}$ and d$\mathbf{r}$ are related as given above.

It is easy to show that Green's theorem is a planar version of this result.

Note that the result is the same for any surface $\mathscr{S}$ whose boundary is $\mathscr{C}$, so any two surfaces $\mathscr{S}_1$, $\mathscr{S}_2$ with the same bounding curve $\mathscr{C}$ give the same surface integral. (We will not give a formal proof of this, but in a nutshell it is because
$\nabla \cdot (\nabla \times \mathbf{F}) = 0$ from Eq. 3.12, then applying the Divergence theorem to the volume enclosed between the two surfaces). This can simplify integration a lot if the bounding curve lies in a plane, since we can replace a surface integral over a curved surface with that over the flat surface with the same boundary.

To emphasize the need for differentiability conditions, consider

$$\mathbf{F} = \frac{-y\mathbf{i} + x\mathbf{j}}{x^2 + y^2}.$$

We can easily verify that $\nabla \times \mathbf{F} = \mathbf{0}$ (except on the z axis where it diverges). But we can also show that $\oint \mathbf{F}.\mathrm{d}\mathbf{r} \neq 0$ if we go around the z axis: for example going round a circle of radius a using a parametrization $(a\cos\theta, a\sin\theta)$we would have

$$\int \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \oint a^{-2}(-a\sin\theta\mathbf{i} + a\cos\theta\mathbf{j}).(-a\sin\theta\mathbf{i} + a\cos\theta\mathbf{j})\,\mathrm{d}\theta = \oint \mathrm{d}\theta = 2\pi.$$

This occurs because our closed curve has looped around the $z-$axis where there is infinite curl; if you do the Complex Variables module in Semester B, this is very similar to a contour integral around a pole.

Example 4.10. Use the surface integral in Stokes's theorem to calculate the circulation of the field $\mathbf{F}$

$$\mathbf{F} = x^2\mathbf{i} + 2x\mathbf{j} + z^2\mathbf{k}$$

around the curve $\mathscr{C}$, where $\mathscr{C}$ is the ellipse $4x^2 + y^2 = 4$ in the x-y plane, taken counterclockwise when viewed from $z > 0$.

In Stokes's Theorem, we can choose **any** surface that spans the curve $\mathscr{C}$. The easiest one in this case is just the planar surface $z = 0$ contained inside the ellipse (so we can use Green's theorem in fact). Thus $\mathbf{n}$ will be purely in the z-direction: $\mathbf{n} = \mathbf{k}$, and so we only need to calculate the z-component of $\nabla \times \mathbf{F}$:

$$(\nabla \times \mathbf{F}).\mathbf{k} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = \frac{\partial(2x)}{\partial x} - \frac{\partial x^2}{\partial y} = 2.$$

Integrating this over the elliptical area is easy: the answer is just 2 times the area of the ellipse. The area of an ellipse is πab, where a is one semi-major axis length (in this case 1) and b is the other semi-major axis length (in this case 2). Hence the answer is 4π.

As in the case of the Divergence Theorem, we can give a vector form of Stokes's Theorem. Given a scalar field U, we let $\mathbf{F} = U\mathbf{a}$ for some constant vector $\mathbf{a}$. Then

$$\begin{aligned}
\int_{\mathscr{C}} U\mathbf{a} \cdot \mathrm{d}\mathbf{r} &= \int_{\mathscr{S}} (\nabla \times (U\mathbf{a})) \cdot d\mathbf{S} \\
&= \int_{\mathscr{S}} ((\nabla U) \times \mathbf{a}) \cdot \mathrm{d}\mathbf{S} \\
&= \mathbf{a}.\int_{\mathscr{S}} \mathrm{d}\mathbf{S} \times (\nabla U).
\end{aligned}$$

The first line is Stokes' theorem, the second follows from the rule Eq 3.8 for curl of a product, and the third from the rules for the scalar triple product. Now we can take the constant $\mathbf{a}$ outside the integral sign; then

choosing **a** = **i**, **j** and **k** in turn, we derive the vector equation

$$\int_{\mathscr{C}} U \, \mathrm{d}\mathbf{r} = \int_{\mathscr{S}} \mathrm{d}\mathbf{S} \times (\nabla U) \qquad .$$

Exercise 4.6. State Stokes's theorem.
Evaluate both sides of the theorem for the vector field $\mathbf{F} = y\mathbf{i} + z\mathbf{j} + y\mathbf{k}$ and the surface S of the hemisphere $x^2 + y^2 + z^2 = 4$ in $z \geq 0$, with normal in the positive z-direction. [You may find the expressions relating Cartesian and spherical polar coordinates useful.] □

Exercise 4.7. Use the surface integral in Stokes's theorem to calculate the circulation of the field **F**

$$\mathbf{F} = 2y\mathbf{i} + 3x\mathbf{j} - z^2\mathbf{k}$$

around the curve $\mathscr{C}$ where $\mathscr{C}$ is the circle $x^2 + y^2 = 9$ in the x-y plane, counterclockwise when viewed from $z > 0$. [Answer 9π.] □

We can use the Divergence and Stokes's theorems to derive other results including, later on, the forms of divergence and curl in curvilinear coordinates in Chapter 5. Those formulas could be found, more laboriously, by direct calculation from the Cartesian definitions by applying the chain rule. Another important application will be given next.

4.7 Conservative Fields and Scalar Potentials

(See Thomas 16.3)
Conservative vector fields play an important role in many applications. A vector field **F** is said to be a **conservative field** iff the value of the line integral $\int_P^Q \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ between endpoints P and Q depends only on the endpoints P and Q, and **not** on the path taken between them. An example of a vector field which is not conservative is the one in Example 4.2 – we explicitly found different answers for the same endpoints, depending on the path taken.

For a conservative vector field **F**, the integral $\int \mathbf{F}.\mathrm{d}\mathbf{r}$ around any **closed** path must be zero (because the value will be given by the trivial path which always stays at the given point). So if **F** is a force, for example, the net work in going round a path back to where one started is zero: energy is conserved, hence the name conservative (nothing to do with politics).

We first state and prove the important result that (subject to differentiability conditions) a vector field is conservative iff it is irrotational (or curl-free). In its statement, 'contractible' means we can continuously deform the region so it squashes to a point. (A torus, for example, is not contractible.)

Theorem 4.4 *In a contractible region,*

$$\nabla \times \mathbf{F} = \mathbf{0} \quad \Longleftrightarrow \quad \exists \text{a scalar field } \phi(\mathbf{r}) \text{ such that } \mathbf{F} = \nabla\phi. \tag{4.12}$$

Note: Such a ϕ is called a **(scalar) potential** for **F**. The theorem says a vector field is conservative iff it has a scalar potential.

Proof:

($\Leftarrow$): This was done at the end of Chapter 3, where we proved the identity $\nabla \times (\nabla\phi) = \mathbf{0}$ for any ϕ, subject to the partial derivatives being well-behaved. Thus if $\mathbf{F} = \nabla\phi$ then $\nabla \times \mathbf{F} = \mathbf{0}$.

($\Rightarrow$): Given $\nabla \times \mathbf{F} = \mathbf{0}$, we proceed by defining the scalar field $\phi(\mathbf{r})$ by

$$\phi(\mathbf{r}) = \int_{\mathbf{a}}^{\mathbf{r}} \mathbf{F}.\mathrm{d}\mathbf{r}, \tag{4.13}$$

where $\mathbf{a}$ is an arbitrary but fixed point; note the line integral has a scalar answer, so ϕ is a scalar field. We will soon show that $\nabla\phi = \mathbf{F}$ as required. First though, since we have not defined the path to be taken from $\mathbf{a}$ to $\mathbf{r}$, we must show that the integral is independent of the path taken, i.e. that the ϕ defined above is well-defined.

Suppose that $\mathscr{C}_1$ and $\mathscr{C}_2$ are two different curves from $\mathbf{a}$ to $\mathbf{r}$. We need to show that

$$\int_{\mathscr{C}_1} \mathbf{F}.\mathrm{d}\mathbf{r} = \int_{\mathscr{C}_2} \mathbf{F}.\mathrm{d}\mathbf{r}.$$

To prove this, let $\mathscr{C}$ be the closed curve formed by following $\mathscr{C}_1$ from $\mathbf{a}$ to $\mathbf{r}$ and then taking $\mathscr{C}_2$ **backwards** to get from $\mathbf{r}$ back to $\mathbf{a}$. Let $\mathscr{S}$ be a surface whose boundary is $\mathscr{C}$. Then:

$$\begin{aligned} \int_{\mathscr{C}_1} \mathbf{F}.\mathrm{d}\mathbf{r} - \int_{\mathscr{C}_2} \mathbf{F}.\mathrm{d}\mathbf{r} &= \int_{\mathscr{C}} \mathbf{F}.\mathrm{d}\mathbf{r} \\ &= \int_{\mathscr{S}} (\nabla \times \mathbf{F}).\mathrm{d}\mathbf{S} \\ &= 0 \end{aligned}$$

The first line is because following $\mathscr{C}_{\in}$ backwards gives us a minus sign in the line integral; the second line is Stokes' theorem for the closed curve $\mathscr{C}$. Hence, the value of ϕ only depends on $\mathbf{r}$, but **not** on the path taken from $\mathbf{a}$ to $\mathbf{r}$, and so $\phi(\mathbf{r})$ **is** well-defined. [Note: Thomas gives a direct proof of the path-independence property for $\mathbf{F} = \nabla V$.]

Next we need to show $\nabla\phi = \mathbf{F}$ as we wanted: we consider a small change $\delta\mathbf{r}$, and we get a small change $\delta\phi$,

$$\delta\phi \equiv \phi(\mathbf{r}+\delta\mathbf{r}) - \phi(\mathbf{r}) = \int_{\mathbf{r}}^{\mathbf{r}+\delta\mathbf{r}} \mathbf{F}\cdot\mathrm{d}\mathbf{r} \approx \mathbf{F}(\mathbf{r})\cdot\delta\mathbf{r}\ ,$$

and this is true for any (infinitesimal) vector $\delta\mathbf{r}$. But by definition of $\nabla\phi$ in Chapter 1, $\delta\phi = (\nabla\phi)\cdot\delta\mathbf{r}$. Hence

$$\nabla\phi\cdot\delta\mathbf{r} = \mathbf{F}\cdot\delta\mathbf{r}\ .$$

But this is true for *all* $\delta\mathbf{r}$, so $\nabla\phi = \mathbf{F}$, as we wanted to show. Q.E.D.

Once we have done this, we easily get the line integral $\int \mathbf{F}\cdot d\mathbf{r}$ between any two points, say $\mathbf{r}_1$ to $\mathbf{r}_2$: choose a path from $\mathbf{r_1}$ back to $\mathbf{a}$, and then from $\mathbf{a}$ to $\mathbf{r_2}$; since taking a line integral backwards gives us a minus sign in the result (as for swapping upper/lower limits in a 1D integral), we get

$$\int_{\mathbf{r_1}}^{\mathbf{r_2}} \mathbf{F}\cdot d\mathbf{r} = \phi(\mathbf{r_2}) - \phi(\mathbf{r_1})$$

Also note that we can add a constant to ϕ without changing $\nabla\phi$; adding a constant is essentially equivalent to changing our choice of fixed point $\mathbf{a}$ in eq. 4.13, since $\phi(\mathbf{a}) = 0$ from the original definition.

In the case where $\mathbf{F}$ is a force, it is usual to define $\phi(\mathbf{r}) = -\int_{\mathbf{a}}^{\mathbf{r}} \mathbf{F}.\mathrm{d}\mathbf{r}$ with an extra (arbitrary) minus sign compared to (4.13); then we get $\mathbf{F} = -\nabla\phi$, and ϕ can then be identified with the potential energy, which decreases when a body moves in the direction of the force "down", and increases in the opposite direction "up". Note again that the value of ϕ is only fixed up to an additive constant, which depends on the choice of reference point $\mathbf{a}$.

Warning: There is a possible snag with notation here: it is very common for historical reasons to use the symbol ϕ (the Greek letter "phi") for a scalar potential, or sometimes V by analogy with Voltage in

electrostatics. That ϕ is obviously **not** related to the coordinate angle ϕ which will appear later in spherical polar coordinates; or also V can possibly get confused with volume. Sometimes the symbols Φ (uppercase phi) or φ (curly phi) are used for the potential, but this still looks quite similar.

Unfortunately, this somewhat confusing notation is heavily used in many textbooks and old exam questions, so it can't be escaped and you just have to be aware of it. In most cases it is reasonably obvious from the context which is which.

Example 4.11. Show that $\mathbf{F} = (z, z, x+y)$ satisfies $\nabla \times \mathbf{F} = 0$, and find a scalar field ϕ such that $\mathbf{F} = \nabla\phi$.

[Note that in answering questions of this sort, where you have to find ϕ, you might as well do that first since $\mathbf{F} = \nabla\phi$ immediately implies $\nabla \times \mathbf{F} = 0$.]

A simple way to do these problems is by direct evaluation of the line integral (4.13), taking as the curve $\mathscr{C}$ as the straight line from the origin (so we are taking **a** to be the origin) to the desired point, (X, Y, Z) say. The line is $\mathbf{r} = t(X, Y, Z)$, $0 \le t \le 1$, so $\mathrm{d}\mathbf{r} = (X, Y, Z)\mathrm{d}t$, while for this example, on that line $\mathbf{F} = (Zt, Zt, Xt+Yt)$. Thus the integral is

$$\int_{\mathscr{C}} \mathbf{F}.\mathrm{d}\mathbf{r} = \int_0^1 [XZt + YZt + (Xt+Yt)Z]\mathrm{d}t = (2XZ+2YZ)\int_0^1 t\,\mathrm{d}t = (2XZ+2YZ)[t^2]_0^1 = XZ+YZ.$$

Hence for a general point we have $\phi = xz + yz$. We can also add any constant to ϕ (since it will disappear in $\nabla\phi$): this expresses the freedom of choice of the **a** in (4.13). [In physical uses of scalar potentials, the reference point is often taken to be at infinity.]

An alternative method is as follows: it is included to emphasize some useful points about integrating sets of partial differential equations (i.e. differential equations with partial derivatives).

We want

$$(z, z, x+y) = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}\right). \tag{4.14}$$

Equating the first components and integrating with respect to x gives

$$z = \frac{\partial\phi}{\partial x} = z \Rightarrow \phi = xz + f(y,z) \tag{4.15}$$

where f is an (as yet) arbitrary function of y and z. Note that f is a 'constant of integration' as far as differentiation with respect to x is concerned: when integrating partial derivatives we have to replace simple constants by functions of those variables not yet taken into account. The second components give

$$z = \frac{\partial\phi}{\partial y} \text{ from (4.14)} = \frac{\partial f}{\partial y} \text{ from (4.15)}.$$

Hence

$$\frac{\partial f}{\partial y} = z \Rightarrow f(y,z) = yz + g(z)\,.$$

No x appears in g since we already know that f does not depend on x. So, substituting this in (4.15),

$$\phi = xz + yz + g(z) \tag{4.16}$$

(g arbitrary as yet). Finally, the third components similarly give

$$x+y = \frac{\partial\phi}{\partial z} \text{ from (4.14)} = x+y+\frac{\mathrm{d}g}{\mathrm{d}z} \text{ from (4.16)}.$$

Hence g has a zero derivative, i.e. is constant and there is a ϕ given by

$$\phi = xz + yz + \text{const.}$$

(We could drop the constant here as without it ϕ would still fulfil the conditions of the problem.) Hence $\nabla \times \mathbf{F} = \mathbf{0}$.

Example 4.12. The gravitational force on a ball of mass m is $\mathbf{F} = (0, 0, -mg)$. If the gravitational acceleration g can be assumed to be constant (which is an excellent approximation for everyday life: $g \simeq 9.8\text{ms}^{-2}$) then $\mathbf{F} = -\nabla\phi$ where $\phi = mgz + \text{const.}$, z being measured, say, from the surface of the Earth. (We can measure z from wherever we wish, since a change of origin just changes the arbitrary constant in ϕ). In this case ϕ is the **gravitational potential energy**.

Exercise 4.8. Show that $\mathbf{F} = (yz, zx, xy)$ is conservative and find a suitable potential ϕ such that $\mathbf{F} = \nabla\phi$. [Answer: $\phi = xyz + \text{const.}$] □

Exercise 4.9. For each of the following fields $\mathbf{F}$, evaluate $\nabla \times \mathbf{F}$ and either find the general solution ϕ satisfying $\mathbf{F} = \nabla\phi$ everywhere, or show that no such ϕ exists:

(a) $\mathbf{F} = x^2\mathbf{i} + y^2\mathbf{j} + 2z\mathbf{k}$

(b) $\mathbf{F} = z^2\mathbf{i} + x^2\mathbf{j} + y^2\mathbf{k}$

(c) $\mathbf{F} = 3z^2\mathbf{i} + 3y^2\mathbf{j} + 6xz\mathbf{k}$

(d) $\mathbf{F} = yz\mathbf{j} - xy\mathbf{k}$.

□

The rest of this chapter will not be lectured and is not examinable. It is included for reference, for completeness, and to give intellectual respectability by proving the main theorems.

4.8 Vector Potentials

(Note: this is not on the syllabus. It is included for completeness, for the sake of those who take later courses where it is used.)

We have seen that, if $\nabla \times \mathbf{F} = \mathbf{0}$, then there exists a scalar potential ϕ such that $\mathbf{F} = \nabla\phi$. There is a similar result if $\nabla \cdot \mathbf{F} = 0$ instead:

Theorem 4.5 *In a contractible domain,*

$$\nabla \cdot \mathbf{F} = 0 \qquad \Longleftrightarrow \qquad \exists \mathbf{A}(\mathbf{r}) \text{ such that } \mathbf{F} = \nabla \times \mathbf{A}.$$

In the $(\Leftarrow)$ direction, this is the identity discussed before. The proof in the other direction consists of writing down suitable integrals, in a way analogous to the proof of (4.12), and is messy so we omit it.

The function $\mathbf{A}$ is called a **vector potential**. Note that one can always add an arbitrary function of the form $\nabla\phi$ to $\mathbf{A}$ and get another perfectly good vector potential for $\mathbf{F}$, because $\nabla \times (\nabla\phi)$ is zero for any ϕ, and so

$$\nabla \times (\mathbf{A} + \nabla\phi) = \nabla \times \mathbf{A} + \nabla \times (\nabla\phi) = \mathbf{F} + \mathbf{0} = \mathbf{F}.$$

In physical contexts this is referred to as a gauge transformation, and provides the basic example whose generalization gives all the modern gauge field theories of physics, the basis of our understanding of all microsopic physical processes.

Example 4.13. Any magnetic field $\mathbf{B}$ satisfies $\nabla \cdot \mathbf{B} = 0$. So, for example, consider a constant magnetic field $\mathbf{B} = (0,0,B_0)$ in the $z-$direction. A suitable vector potential $\mathbf{A}$ in this case is

$$\left(-\frac{1}{2}B_0 y, \frac{1}{2}B_0 x, 0\right),$$

since

$$\begin{aligned}\nabla \times \mathbf{A} &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}, \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) \\ &= \left(0-0, 0-0, \frac{1}{2}B_0 - (-\frac{1}{2}B_0)\right) \\ &= \mathbf{B}.\end{aligned}$$

4.9 Derivations of the main theorems

(See Thomas 16.7 and 16.8)
[This section is not examinable]

We now return to the proofs of the Divergence and Stokes's Theorems.

Consider first the "proof" of the Divergence Theorem using rectangular boxes. Take a box $[x_1, x_2] \times [y_1, y_2] \times [z_1, z_2]$. Then for a vector $\mathbf{A} = A_1\mathbf{i} + A_2\mathbf{j} + A_3\mathbf{k}$,

$$\begin{aligned}\int (\nabla \cdot \mathbf{A}) \mathrm{d}V &= \int\int\int \left(\frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} + \frac{\partial A_3}{\partial z}\right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ &= \int\int [A_1]_{x_1}^{x_2} \mathrm{d}y\,\mathrm{d}z + \int\int [A_2]_{y_1}^{y_2} \mathrm{d}x\,\mathrm{d}z + \int\int [A_3]_{z_1}^{z_2} \mathrm{d}x\,\mathrm{d}y \\ &= \int\int_{\text{front}} A_1 \mathrm{d}y\,\mathrm{d}z - \int\int_{\text{back}} A_1 \mathrm{d}y\,\mathrm{d}z + \int\int_{\text{right end}} A_2 \mathrm{d}x\,\mathrm{d}z - \int\int_{\text{left end}} A_2 \mathrm{d}x\,\mathrm{d}z \quad (4.17) \\ &\quad + \int\int_{\text{top}} A_3 \mathrm{d}x\,\mathrm{d}y - \int\int_{\text{bottom}} A_3 \mathrm{d}x\,\mathrm{d}y.\end{aligned}$$

On the front of the box (i.e. the surface $x = x_2$) $\mathrm{d}\mathbf{S} = \mathbf{i}\mathrm{d}y\mathrm{d}z$ while on the back ($x = x_1$) $\mathrm{d}\mathbf{S} = -\mathbf{i}\mathrm{d}y\mathrm{d}z$ so the first two terms in (4.17) are $\int \mathbf{A}.\mathrm{d}\mathbf{S}$ for the front and back. Similarly for the remaining terms.

One can complete a "proof" by decomposing a volume into such boxes and adding the results, noting that the surface integrals on a face common to two boxes will cancel one another. This overlooks the difficulty of proving that the surface integral for all the boxes gives a correct limit for the smooth surface (for the volume integral this just follows from the definition of such integrals).

Instead we can work towards a correct proof by first noting that the terms match up in the sense that

$$\int\int\int_{\mathscr{D}} \frac{\partial A_3}{\partial z} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int\int_{\mathscr{S}} A_3 (\mathrm{d}S)_z \qquad (4.18)$$

for the box. (What we thus really do is prove the theorem for $\mathbf{F} = A_3\mathbf{k}$ and then add together three such results.)

We now have to cope with some technical points
1. We must be able to integrate the derivatives of **A** once. A sufficient condition is that all first derivatives of **A** are piecewise continuous. If the derivatives have discontinuities we have to do the proof for each smooth piece separately and then add the results.
2. That first point implies **A** itself must be piecewise continuous.
3. We require the surface to be bounded (so we have a finite area) and closed (so we have a finite volume).
4. We must to be able to integrate $\int \mathbf{A}.\mathrm{d}\mathbf{S}$. So we want to be able to assign coordinates on pieces of the surface S, say (u, v), in such a way that $(\mathbf{e}_u \times \mathbf{e}_v)\mathrm{d}u\,\mathrm{d}v$ can be defined and calculated, i.e. we want the map $\mathbb{R}^2 \to \mathbb{R}^3 : (u, v) \to (x(u, v), y(u, v), z(u, v))$ to be (piecewise) sufficiently differentiable.

These assumptions ensure we can break D up into convex pieces. 'Convex' means that any line cuts the surface at most twice. So now we have the form

Theorem 4.6 *If $\mathscr{S}$ is a bounded closed piecewise smooth orientable surface enclosing a volume $\mathscr{D}$, and if* **F** *is a vector field all of whose first derivatives are continuous, then*

$$\int_{\mathscr{D}} \nabla \cdot \mathbf{F}\,\mathrm{d}V = \int_{\mathscr{S}} \mathbf{F}.\mathbf{n}\,\mathrm{d}S = \int_{\mathscr{S}} \mathbf{F}.\mathrm{d}\mathbf{S},$$

where **n** *is the normal outward-pointing from $\mathscr{D}$.*

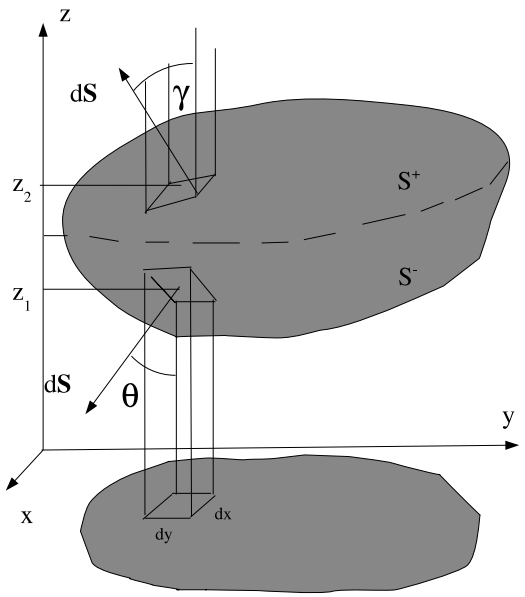

Figure 4.3: Convex surface used in the proof of the Divergence Theorem

Proof: [This proof is more-or-less identical, with slight changes in notation, with the one given by Thomas.] We break $\mathscr{D}$ into convex pieces and first prove the result for a single convex piece (which we call $\mathscr{D}_1$). In fact we need only prove (4.18). Consider lines parallel to the z-axis. Those which meet $\mathscr{D}_1$ either meet it twice or touch it on a closed curve. Divide the surface into $\mathscr{S}^+$ and $\mathscr{S}^-$, the upper and lower halves (i.e. $\mathscr{S}^-$ is where the lines parallel to the z-axis first meet $\mathscr{S}$: see Figure 4.3). Then, just using the fundamental theorem of calculus,

$$\int\int\int_{\mathscr{D}_1} \frac{\partial A_3}{\partial z}\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int\int_{\mathscr{S}^+} A_3(x, y, z_2)\mathrm{d}x\,\mathrm{d}y - \int\int_{\mathscr{S}^-} A_3(x, y, z_1)\mathrm{d}x\,\mathrm{d}y$$

On $\mathscr{S}^+$, $(A_3\mathbf{k}).\mathrm{d}\mathbf{S} = A_3|\mathrm{d}\mathbf{S}|\cos\gamma = A_3\mathrm{d}x\,\mathrm{d}y$ and similarly on $\mathscr{S}^-$. Hence we have shown that

$$\int\int\int_{\mathscr{D}_1} \frac{\partial A_3}{\partial z}\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int\int_{\mathscr{S}} (A_3\mathbf{k}).\mathrm{d}\mathbf{S}.$$

and adding similar results for A_1 and A_2 we get the Divergence Theorem for $\mathscr{D}_1$. When we re-combine the convex pieces, the surfaces where they join appear twice in the surface integrals, once with each of the two possible signs for the normal, so these parts cancel one another and only the integral over the bounding surface remains. Q.E.D.

We showed above that the Divergence Theorem implies Green's theorem. We only have Stokes's theorem left to prove. The conditions are arrived at by similar considerations to those for the Divergence Theorem.

Theorem 4.7 *For any piecewise smooth surface $\mathscr{S}$ bounded by a piecewise smooth curve $\mathscr{C}$ on which $\nabla\times\mathbf{F}$ is piecewise continuous,*

$$\int_{\mathscr{S}} \nabla\times\mathbf{F}.\mathrm{d}\mathbf{S} = \oint_{\mathscr{C}} \mathbf{F}.\mathrm{d}\mathbf{r},$$

where the integral round $\mathscr{C}$ is taken in the direction which is counter-clockwise as seen from the side of $\mathscr{S}$ pointed to by d**S**.

Proof: The conditions imply that the surface can be decomposed in pieces which project to regions in one of the planes of Cartesian coordinates; without loss of generality say the (x,y) plane. We prove the result for one such region. Suppose we have coordinates (u,v) on this region. We also consider only the terms involving P where $\mathbf{F} = (P,Q,R)$ (i.e. we prove the result for $\mathbf{F} = P\mathbf{i}$ first).

$$\begin{aligned}
\oint_{\mathscr{C}} P\mathrm{d}x &= \oint_{\mathscr{C}} P\left(\frac{\partial x}{\partial u}\mathrm{d}u + \frac{\partial x}{\partial v}\mathrm{d}v\right)\\
&= \int\int\left[-\frac{\partial}{\partial v}\left(P\left(\frac{\partial x}{\partial u}\right)\right) + \frac{\partial}{\partial u}\left(P\left(\frac{\partial x}{\partial v}\right)\right)\right]\mathrm{d}u\,\mathrm{d}v \text{ by Green's theorem}\\
&= \int\int\left(\frac{\partial P}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial P}{\partial v}\frac{\partial x}{\partial u}\right)\mathrm{d}u\,\mathrm{d}v\\
&= \int\int\left(\left(\frac{\partial P}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial P}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial u}\right)\frac{\partial x}{\partial v} - \left(\frac{\partial P}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial P}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial v}\right)\frac{\partial x}{\partial u}\right)\mathrm{d}u\,\mathrm{d}v\\
&\qquad \text{using the Chain Rule}\\
&= \int\int\frac{\partial P}{\partial y}\left(\frac{\partial y}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial y}{\partial v}\frac{\partial x}{\partial u}\right)\mathrm{d}u\,\mathrm{d}v + \int\int\frac{\partial P}{\partial z}\left(\frac{\partial z}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial z}{\partial v}\frac{\partial x}{\partial u}\right)\mathrm{d}u\,\mathrm{d}v
\end{aligned}$$

and taking the cross product of

$$\begin{aligned}
\mathrm{d}\mathbf{r}_u &= \left(\frac{\partial x}{\partial u}\mathbf{i} + \frac{\partial y}{\partial u}\mathbf{j} + \frac{\partial z}{\partial u}\mathbf{k}\right)\mathrm{d}u,\\
\mathrm{d}\mathbf{r}_v &= \left(\frac{\partial x}{\partial v}\mathbf{i} + \frac{\partial y}{\partial v}\mathbf{j} + \frac{\partial z}{\partial v}\mathbf{k}\right)\mathrm{d}v,
\end{aligned}$$

easily shows that the double integrals give

$$\int\int\left(-\frac{\partial P}{\partial y}(\mathrm{d}\mathbf{S})_z + \frac{\partial P}{\partial z}(\mathrm{d}\mathbf{S})_y\right)$$

which is the part of $\nabla\times\mathbf{F}.\mathrm{d}\mathbf{S}$ involving P. To complete the proof we add the parts with Q and R and add together the results from the pieces into which a general $\mathscr{S}$ has to be split. Q.E.D.

Chapter 5

Orthogonal Curvilinear Coordinates

Last update: 22 Nov 2010

Syllabus section:
4. Orthogonal curvilinear coordinates; length of line element; grad, div and curl in curvilinear coordinates; spherical and cylindrical polar coordinates as examples.

So far we have only used Cartesian x,y,z coordinates. Sometimes, because of the geometry of a given problem, it is easier to work in some other coordinate system. Here we show how to do this, restricting the generality only by an orthogonality condition.

5.1 Plane Polar Coordinates

In Calculus II and Chapter 2, we met the simple curvilinear coordinates in two dimensions, plane polars, defined by

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

We can easily invert these relations to get

$$r = \sqrt{x^2+y^2}, \qquad \theta = \arctan(y/x).$$

The Chain Rule enables us to relate partial derivatives with respect to x and y to those with respect to r and θ and vice versa, e.g.

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} \qquad = \frac{x}{r}\frac{\partial f}{\partial x} + \frac{y}{r}\frac{\partial f}{\partial y}\ . \tag{5.1}$$

In Calculus II, the rule for changing coordinates in integrals is also given. The general rule is that if we change coordinates from x,y to u,v where $x = x(u,v)$, $y = y(u,v)$, then a $\mathrm{d}x\,\mathrm{d}y$ in an area integral is replaced by the Jacobian determinant

$$\left|\begin{array}{cc} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{array}\right| \mathrm{d}u\,\mathrm{d}v\ .$$

If we define $\mathbf{r} = (x(u,v), y(u,v), 0)$, differentiate w.r.t. u,v and take the cross-product, we will see that the above is equal to

$$\mathrm{d}S = \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| \mathrm{d}u\,\mathrm{d}v \qquad ,$$

as we derived in section 4.2.

For plane polar coordinates, replacing u, v with r, θ and calcuating the determinant above gives $\mathrm{d}S = r\,\mathrm{d}r\,\mathrm{d}\theta$; this can also be shown geometrically by considering an infinitesimal quadrilateral with corners at $(r,\theta),\ldots,(r+dr,\theta+d\theta)$ and working out the area from a sketch.

Example 5.1. The Gaussian integral (related to the Gaussian distribution in statistics)

Consider the integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,dx\,dy = \left(\int_{-\infty}^{\infty} e^{-x^2}\,dx\right)^2.$$

Transforming to polar coordinates gives

$$\int_0^{\infty} re^{-r^2}\,\mathrm{d}r \int_0^{2\pi}\mathrm{d}\theta = [-\tfrac{1}{2}e^{-r^2}]_0^{\infty}[\theta]_0^{2\pi} = \pi$$

and hence (according to Dr. Saha "the most beautiful of all integrals")

$$\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

For later use, we now construct the unit vectors in the directions in which r and θ increase at a point, which we will denote $\mathbf{e}_r$ and $\mathbf{e}_\theta$. These are tangent to the coordinate lines, where a **coordinate line** means a curve on which only one of the coordinates is varying, and the other coordinates are fixed. Coordinate lines are generalizations of lines parallel to the x, y, z axes in Cartesians, but now they won't be straight lines (hence the "**curvilinear**" in the chapter title).

We already know how to find the tangent vectors to coordinate lines, by taking partial derivatives of $\mathbf{r}$ with respect to each of r, θ; then all we have to do is divide those by their lengths to get unit vectors. Thus in plane polars we have

$$\mathbf{r} = r\cos\theta\,\mathbf{i} + r\sin\theta\,\mathbf{j} \quad ,$$

so a small change dr gives us a change

$$\mathrm{d}\mathbf{r}_r = \frac{\partial \mathbf{r}}{\partial r}\mathrm{d}r = (\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j})\mathrm{d}r, \qquad \left|\frac{\partial \mathbf{r}}{\partial r}\right| = 1 \quad \Rightarrow \quad \mathbf{e}_r = \frac{\partial \mathbf{r}}{\partial r} = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}$$

while a small change $d\theta$ gives us

$$\mathrm{d}\mathbf{r}_\theta = \frac{\partial \mathbf{r}}{\partial \theta}\mathrm{d}\theta = (-r\sin\theta\,\mathbf{i} + r\cos\theta\,\mathbf{j})\mathrm{d}\theta, \qquad \left|\frac{\partial \mathbf{r}}{\partial \theta}\right| = r \quad \Rightarrow \quad \mathbf{e}_\theta = -\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j}.$$

So a general small displacement becomes

$$\delta\mathbf{r} = \mathbf{e}_r\,\delta r + r\,\mathbf{e}_\theta\,\delta\theta \quad .$$

We will see the value of this later on; we are next going to consider three-dimensional versions of polar coordinates: there are two common versions, firstly cylindrical polars and then spherical polars.

5.2 Cylindrical Polar Coordinates

For cylindrical polars, we turn the plane polars in the x, y plane into three-dimensional coordinates by simply using z as the third coordinate (see Fig. 5.1). To avoid confusion with other coordinate systems, we shall for clarity [1] rename r as ρ and θ as ϕ , but beware that in other courses, books, and applications of these ideas, r and θ will still be used. Thus we have

$$x = \rho\cos\phi, \quad y = \rho\sin\phi, \quad z = z\ ,$$

or

$$\mathbf{r} = \rho\cos\phi\,\mathbf{i} + \rho\sin\phi\,\mathbf{j} + z\mathbf{k} \quad ,$$

and quantities in any plane z =constant will be as in plane polars. The figure 5.1 shows coordinate lines for each of ρ, ϕ and z; here the coordinate line for ρ is a line of varying ρ and constant ϕ,z; and likewise for the other two. Note that the coordinate lines for ρ, z are straight lines, while the ϕ line is a circle around the z axis. Thomas's Fig. 15.37 shows a nice diagram of surfaces on which one of the coordinates is constant: the constant-ρ surface is a cylinder whose axis is the $z-$axis, while surfaces of constant ϕ or constant z are planes.

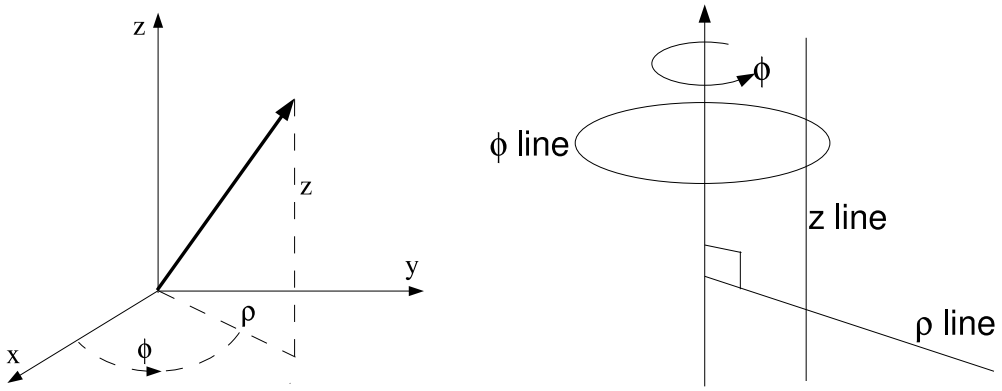

Figure 5.1: Cylindrical polar coordinates relative to Cartesian, and with sample ρ- and ϕ-curves shown.

The fact that constant ρ gives a cylinder gives the name cylindrical polars: these coordinates are natural ones to use whenever there is a problem involving cylindrical geometry or symmetry (for example, doing a surface integration over a cylinder, or in physics calculating a magnetic field around a straight wire).

To get partial derivatives in curvilinear coordinates we again use the chain rule (5.1), but now with three terms on the right. Taking the plane polar results, changing variable names and appending $\mathbf{e}_z = \mathbf{k}$, the unit vectors along the coordinate lines are

$$\mathbf{e}_\rho = \cos\phi\,\mathbf{i} + \sin\phi\,\mathbf{j}\ , \quad \mathbf{e}_\phi = -\sin\phi\,\mathbf{i} + \cos\phi\,\mathbf{j}\ , \quad \mathbf{e}_z = \mathbf{k}$$

respectively. We can write this in matrix form as

$$\begin{pmatrix} \mathbf{e}_\rho \\ \mathbf{e}_\phi \\ \mathbf{e}_z \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix}. \tag{5.2}$$

[1] Unfortunately, for the same reasons of clarity, Thomas adopts the alternative solution of renaming two of the spherical polar coordinates. To avoid confusion with past years' exam papers I have kept to the choice used there, which is also the one used in most books. Thomas chooses (ρ, ϕ, θ) for the usual (r, θ, ϕ). The swap of θ and ϕ is particularly likely to be confusing.

It is easy to see from the above that the dot-product of any two **e**'s gives 1 (if they are the same) or 0 (for any two different ones), like the rules for $\mathbf{i},\mathbf{j},\mathbf{k}$. This implies that the three **e**'s are an **orthogonal triple** of unit vectors, and also implies geometrically that the cross-product of any two different **e**'s will be $\pm$ the third one.

We can also express this property in matrix notation: the 3×3 matrix above, call it $\mathbf{R}$, is a rotation matrix, i.e. one such that $\mathbf{R}^{-1}=\mathbf{R}^T$, where the $\mathbf{R}^T$ denotes transpose. This comes about because the dot-product of any two **e**'s is given by one element of the matrix $\mathbf{R}\mathbf{R}^T$, and the **e**'s are an orthogonal triple if and only if $\mathbf{R}\mathbf{R}^T=\mathbf{I}$, the identity matrix. [2] Also note that if we want $\mathbf{i},\mathbf{j},\mathbf{k}$ in terms of the **e**'s, we can just multiply Eq. 5.2 by $\mathbf{R}^{-1}=\mathbf{R}^T$.

The lengths of $\partial\mathbf{r}/\partial\rho$, $\partial\mathbf{r}/\partial\phi$ and $\partial\mathbf{r}/\partial z$ are respectively 1, ρ and 1; we can use these together with the **e**'s to find infinitesimal area elements: e.g. taking a surface $\rho=$constant (a cylinder), we can treat this as a 2-parameter surface with ϕ,z as the parameters, so the vector area element for small changes $d\phi,dz$ is given by

$$\begin{aligned} \mathrm{d}\mathbf{S} &= \frac{\partial\mathbf{r}}{\partial\phi}\times\frac{\partial\mathbf{r}}{\partial z}\,\mathrm{d}\phi\,\mathrm{d}z \\ &= \rho\mathbf{e}_\phi\times\mathbf{e}_z\,\mathrm{d}\phi\,\mathrm{d}z \\ &= \rho\ \mathbf{e}_\rho\mathrm{d}\phi\,\mathrm{d}z\ ; \end{aligned}$$

this will be useful when doing surface integrals over a cylinder. (As usual, there is a potentially ambiguous choice of sign with vector areas, due to the sign-flip in changing order of a cross product; take care with this, e.g. when doing a problem check that your vector area matches the desired direction).

When doing volume integrals, we may need the volume element which is

$$\mathrm{d}V=\rho\ \mathrm{d}\rho\,\mathrm{d}\phi\mathrm{d}z$$

from the scalar triple product.

5.3 Spherical Polar Coordinates

These are coordinates (r,θ,ϕ), where r measures distance from the origin, θ measures angle from some chosen axis, called the **polar axis**, and ϕ measures angle around that axis (see Fig 5.2.) To relate them to Cartesian coordinates we usually assume that the z-axis is the polar axis. Then, let P be our chosen point at (r,θ,ϕ), and drop a perpendicular from P to the $z-$axis meeting it at Q. The line OP is at angle θ to the positive z-axis, so clearly $OQ=z=r\cos\theta$ and PQ = $r\sin\theta$. Dropping another perpendicular from P to the *xy* plane, we get a point in the *xy* plane at distance $r\sin\theta$ from the origin; then inserting $\rho=r\sin\theta$ into the cylindrical polars in Sec. 5.2 gives us:

$$x=r\sin\theta\cos\phi,\quad y=r\sin\theta\sin\phi,\quad z=r\cos\theta.$$

or, as a position vector

$$\mathbf{r}=r\sin\theta\cos\phi\,\mathbf{i}+r\sin\theta\sin\phi\,\mathbf{j}+r\cos\theta\,\mathbf{k}$$

Here the ϕ is the same as that of cylindrical polars, which explains why we chose the same letter. The inverse of these relations is

$$r=\sqrt{x^2+y^2+z^2},\quad \theta=\arctan\left(\frac{\sqrt{x^2+y^2}}{z}\right),\quad \phi=\arctan\left(\frac{y}{x}\right).$$

[2] Rotation matrices are "special" because they preserve lengths and angles; e.g. if we take two vectors **a**, **b**, write them as column vectors, then their scalar product in matrix notation is $\mathbf{a}^T\mathbf{b}$. The two vectors rotated by matrix **R** are **Ra** and **Rb**. To conserve scalar product, we must have $(\mathbf{R}\mathbf{a})^T(\mathbf{R}\mathbf{b})=\mathbf{a}^T\mathbf{b}$, and using the transpose rule this becomes $\mathbf{a}^T\mathbf{R}^T\mathbf{R}\mathbf{b}=\mathbf{a}^T\mathbf{b}$. For this to apply for *any* two $\mathbf{a},\mathbf{b}$ we must have $\mathbf{R}^T\mathbf{R}=\mathbf{I}$, the identity matrix.

Coordinate lines of r, (i.e. lines of constant θ and ϕ), are straight radial lines from the origin; coordinate

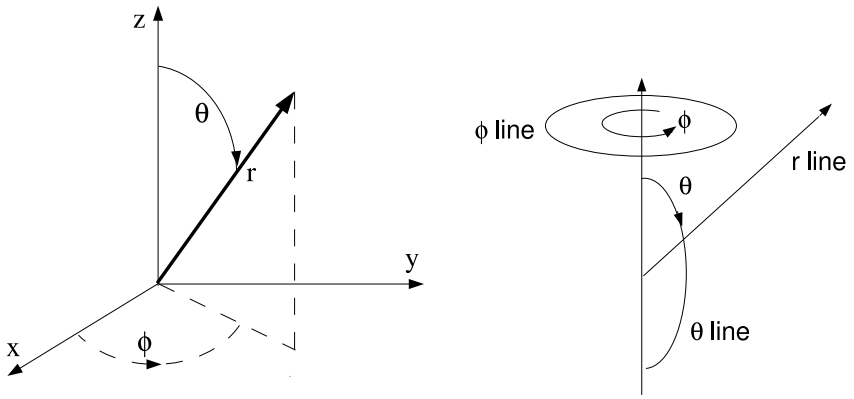

Figure 5.2: Spherical polar coordinates relative to Cartesian, and with sample r-, $\theta-$ and ϕ-curves shown.

lines of θ (constant r and ϕ) are **meridional** semicircles, i.e. semicircles centred at the origin and in a plane containing the polar axis; and coordinate lines of ϕ (constant r and θ) are **latitudinal** circles, i.e. circles centred at a point on the polar axis and in a plane perpendicular to it. Note however that while r runs from 0 to ∞ (like the r of plane polars and ρ of cylindrical polars) and ϕ runs from 0 to 2π (like the θ of plane polars), θ only runs **from 0 to** π, since for any point P the angle between OP and the z-axis won't exceed 180 degrees = π radians.

The coordinate lines of θ are strictly semi-circles, rather than circles. To make a circle we have to take the coordinate lines of θ for two different ϕ, say ϕ_0 and $\phi_0+\pi$. Thomas's Fig. 15.42 shows a nice diagram of surfaces on which one of the coordinates is constant.

You should beware of the fact that some authors, including Thomas, use different notation, in particular swapping the meanings of θ and ϕ in the definition of spherical polars. We shall consistently use the above notation for spherical polar coordinates, which is the most common one, throughout this course.

Note that these again generalize the plane polar coordinates, but this time the polars r, θ are in planes containing the z (or polar) axis, rather than in planes perpendicular to it. The spherical polar coordinates are of course the natural ones to use when we have a spherical geometry, or part of a sphere.

Now we construct the **e** vectors as before: taking partial derivatives of **r** above with respect to each of the coordinates in turn, we get

$$\begin{aligned}
\partial\mathbf{r}/\partial r &= \sin\theta\cos\phi\,\mathbf{i}+\sin\theta\sin\phi\,\mathbf{j}+\cos\theta\,\mathbf{k},\\
\partial\mathbf{r}/\partial\theta &= r\cos\theta\cos\phi\,\mathbf{i}+r\cos\theta\sin\phi\,\mathbf{j}-r\sin\theta\,\mathbf{k}\\
\partial\mathbf{r}/\partial\phi &= -r\sin\theta\sin\phi\,\mathbf{i}+r\sin\theta\cos\phi\,\mathbf{j}.
\end{aligned}$$

The lengths of these, by simple applications of $\cos^2\phi+\sin^2\phi=1$, are respectively 1, r, and $r\sin\theta$. Dividing these derivatives by their lengths gives us the unit vectors $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ tangent to the coordinate lines, which we can write as

$$\begin{pmatrix}\mathbf{e}_r\\ \mathbf{e}_\theta\\ \mathbf{e}_\phi\end{pmatrix}=\begin{pmatrix}\sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta\\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta\\ -\sin\phi & \cos\phi & 0\end{pmatrix}\begin{pmatrix}\mathbf{i}\\ \mathbf{j}\\ \mathbf{k}\end{pmatrix}. \tag{5.3}$$

It is straightforward to show that again the dot-product of any two **e**'s is 1 (if they are the same) or 0 (if different); therefore the cross-product of any two **e**'s is $\pm$ the third one and the matrix above is again a rotation matrix.

It is also worth noting that $\mathbf{e}_r = \mathbf{r}/r$, as expected by symmetry since $\mathbf{e}_r$ is a unit vector pointing away from the origin at point $\mathbf{r}$.

For doing integrals later on, the volume element is given by the scalar triple product

$$dV = (\partial\mathbf{r}/\partial r) \times (\partial\mathbf{r}/\partial\theta).(\partial\mathbf{r}/\partial\phi)\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\phi = r^2 \sin\theta\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\phi \quad .$$

The infinitesimal area element on a sphere (i.e. a surface of constant r) is given by

$$d\mathbf{S} = (\partial\mathbf{r}/\partial\theta) \times (\partial\mathbf{r}/\partial\phi)\, \mathrm{d}\theta\, \mathrm{d}\phi \quad = r^2 \sin\theta\, \mathbf{e}_r\, \mathrm{d}\theta \mathrm{d}\phi \qquad .$$

Similar results hold for surfaces of constant ϕ and of constant θ, but are not so common in practice; note that the above area element on a sphere turns up in many examples and exam questions, and is well worth memorising.

Example 5.2. "Earth polar coordinates"

To define spherical polars on the Earth, let the polar axis be the Earth's rotation axis, with z increasing to the North, let the equator define the x,y plane, and let the prime meridian (the one through Greenwich) be $\phi = 0$. Then any point on the Earth's surface can be referred to by the spherical polar angles (θ, ϕ). In navigation people use latitude and logitude. Longitude is measured East or West from the prime meridian and is in the range $(0, 180^\circ)$ so to get ϕ for a place with Westerly longitude we just subtract from $2\pi = 360^\circ$. Latitude is defined to be 0 at the equator (whereas $\theta = 90^\circ = \pi/2$ there). Given a latitude, we need to subtract it from 90° if it is North and add it to 90° if it is South.

For example Buenos Aires, which has latitude $34^\circ 36'$S, and longitude $58^\circ 22'$W, will have Earth polar coordinates $\theta = 125^\circ, \phi = 302^\circ$ to the nearest degree.

5.4 Some applications of these polar coordinates

Using polar (or cylindrical) coordinates the area within a circle of radius R, $\int_0^R \int_0^{2\pi} r\, \mathrm{d}\phi\, \mathrm{d}r$, comes out immediately as πR^2.

Using spherical polar coordinates the volume of a sphere of radius R is

$$\int_0^R \int_0^\pi \int_0^{2\pi} r^2 \sin\theta\, \mathrm{d}\phi\, \mathrm{d}\theta\, \mathrm{d}r$$

which evaluates to $\frac{4}{3}\pi R^3$. (Remember that for a full sphere, the ranges of integration are $0 \le \theta \le \pi$, $0 \le \phi \le 2\pi$).

Example 5.3. Area of a cone:

Consider the conical surface $\theta = \theta_1$ cut in a sphere of radius s. The area is given by integrating

$$\int_0^{2\pi} \mathrm{d}\phi \int_0^s \sin\theta_1 r\, \mathrm{d}r = \pi s^2 \sin\theta_1 .$$

Here s is the slant height of the cone. The cone's base (say b) will be $s\sin\theta_1$. Hence we can express the sloping area of a cone neatly as πsb.

Example 5.4. We now reconsider Example 4.5.
Find the flux of the field $\mathbf{F} = z\mathbf{k}$ across the portion of the sphere $x^2 + y^2 + z^2 = a^2$ in the first octant with normal taken in the direction away from the origin.

Because of the geometry of the surface, it is easiest to work in spherical polar coordinates (r,θ,ϕ), so the sphere has $r=a$. The unit normal $\mathbf{n}$ to the sphere that points away from the origin is just $\mathbf{e}_r$, the outward radial vector of unit length. Now

$$\mathbf{F}.\mathbf{e}_r = z\mathbf{k}\cdot\mathbf{e}_r = z\cos\theta = r\cos^2\theta\ .$$

using 5.3 to evaluate $\mathbf{k}.\mathbf{e}_r=\cos\theta$. An area element on the surface of a sphere of radius r is $(r\mathrm{d}\theta)(r\sin\theta\,\mathrm{d}\phi) = r^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$. For our given sphere $r=a$, so

$$\begin{aligned}\int_S \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S &= \int_0^{\pi/2}\int_0^{\pi/2} a\cos^2\theta\, a^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \\ &= a^3\int_0^{\pi/2}\int_0^{\pi/2}\cos^2\theta\sin\theta\,\mathrm{d}\theta\mathrm{d}\phi \\ &= \frac{\pi}{2}a^3\left[\frac{-1}{3}\cos^3\theta\right]_0^{\pi/2} \\ &= \frac{\pi}{6}a^3\ .\end{aligned}$$

Note that the integrand didn't depend on ϕ, so we just replaced the $d\phi$ integral with a multiplication by the range, here $(\pi/2-0)$. This is a common short-cut to note.

Example 5.5. Cutting an apple

In his book, Matthews poses a good problem for illustrating integration using curved coordinates: "A cylindrical apple corer of radius a cuts through a spherical apple of radius b. How much of the apple does it remove?"

We can reformulate the problem slightly, without losing generality, by letting the radius of the apple equal unity and introducing $\sin\theta_1 = a/b$ (i.e. we scale the problem by b). In our restated problem the corer cuts through the peel at $\theta=\theta_1$ and $\theta=\frac{1}{2}\pi-\theta_1$ in spherical polars, i.e. in cylindrical polars at

$$\rho=\sin\theta_1,\quad z=\cos\theta_1,$$

and, of course, at $z=-\cos\theta_1$.

We can now complete the solution of this problem in (at least) four different ways: three of these are relegated to an appendix, not given in lectures.[3]

The first way is to integrate over z and then ρ

$$4\pi\int_0^{\sin\phi_1}\rho\,d\rho\int_0^{\sqrt{1-\rho^2}}dz = 4\pi\int_0^{\sin\phi_1}\rho(1-\rho^2)^{\frac{1}{2}}\,d\rho = \frac{4\pi}{3}(1-\cos^3\phi_1).$$

5.5 General Orthogonal Curvilinear Coordinates

The two sets of polar coordinates above have a feature in common: the three sets of coordinate lines are orthogonal to one another at all points, because their tangent vectors and corresponding unit vectors **e**'s are orthogonal. (This is where the **orthogonal** in the chapter title comes from).

[3] I give only the key steps. Some algebraic filling-in is needed. In each version we can shorten the calculations by replacing the ϕ integration with multiplication by 2π (since the integrand doesn't depend on ϕ) , and also doing the integrals only for $z\geq 0$, and then doubling using symmetry.

General orthogonal coordinates are coordinates for which these properties are true, i.e. the coordinate lines are always mutually perpendicular **at a given point**, though they are generally curved. In general, coordinates need not be orthogonal. However, we shall be concerned only with orthogonal curvilinear coordinates. Cylindrical polars and spherical polars are the only non-Cartesian coordinate systems in which you will be expected to perform explicit calculations in this course, apart from simple substitutions into the general formulae.

Suppose (u_1, u_2, u_3) are a general set of coordinates, defined by some given function $\mathbf{r}(u_1, u_2, u_3)$. As before, we calculate $\partial\mathbf{r}/\partial u_1$ which is the tangent vector to a u_1 line (varying u_1, constant u_2, u_3). Next we define the arc-length h_1 and unit vector $\mathbf{e}_1$ as

$$h_1 = \left|\frac{\partial \mathbf{r}}{\partial u_1}\right|, \qquad \mathbf{e}_1 = \frac{\partial \mathbf{r}}{\partial u_1}/h_1$$

therefore

$$\frac{\partial \mathbf{r}}{\partial u_1} \equiv h_1 \mathbf{e}_1 \quad .$$

It is easy to calculate that

$$h_1^2 = \left(\frac{\partial x}{\partial u_1}\right)^2 + \left(\frac{\partial y}{\partial u_1}\right)^2 + \left(\frac{\partial z}{\partial u_1}\right)^2 .$$

Likewise differentiating $\mathbf{r}$ by u_2, u_3, we define two more unit vectors $\mathbf{e}_2$, $\mathbf{e}_3$, along the coordinate lines of u_2 and u_3, and associated arc-length parameters h_2 and h_3. This is useful for several reasons: firstly, $\mathbf{e}_1$ tells us in which direction $\mathbf{r}$ moves with a small change in u_1, while $h_1\, du_1$ is the distance moved along $\mathbf{e}_1$, and likewise for changes du_2, du_3.

We define a coordinate system to be **orthogonal** iff $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ are mutually orthogonal everywhere:

$$\text{Coordinates } (u_1, u_2, u_3) \text{ are orthogonal} \Leftrightarrow \frac{\partial \mathbf{r}}{\partial u_1}.\frac{\partial \mathbf{r}}{\partial u_2} = \frac{\partial \mathbf{r}}{\partial u_2}.\frac{\partial \mathbf{r}}{\partial u_3} = \frac{\partial \mathbf{r}}{\partial u_3}.\frac{\partial \mathbf{r}}{\partial u_1} = 0$$

For orthogonal coordinates, a general small change $(\mathrm{d}u_1, \mathrm{d}u_2, \mathrm{d}u_3)$ in the coordinates means a displacement

$$d\mathbf{r} \quad = \quad h_1 \mathrm{d}u_1 \mathbf{e}_1 + h_2 \mathrm{d}u_2 \mathbf{e}_2 + h_3 \mathrm{d}u_3 \mathbf{e}_3 \quad , \tag{5.4}$$

which corresponds to a distance

$$\left(h_1^2 \mathrm{d}u_1^2 + h_2^2 \mathrm{d}u_2^2 + h_3^2 \mathrm{d}u_3^2\right)^{1/2} .$$

Also, for orthogonal coordinates the dot and cross products of any two **e**'s will obey the same rules we met before: therefore the matrix $\mathbf{R}$ relating $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ to $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ will be a rotation matrix (from above) and have the property that $\mathbf{R}^T = \mathbf{R}^{-1}$.

Cartesian coordinates are of course a special simple case of orthogonal curvilinear coordinates, in which all the coordinate lines are straight lines and all of $h_1 = h_2 = h_3 = 1$.

Sometimes it is convenient to replace the 1,2,3 with the letters of the coordinates, e.g. in cylindrical polar coordinates, we wrote $\mathbf{e}_\rho$, $\mathbf{e}_\phi$, $\mathbf{e}_z$. There we already found $h_\rho = 1$ and $h_z = 1$, but $h_\phi = \rho$, so a change $\mathrm{d}\phi$ corresponds to moving a distance $\rho \mathrm{d}\phi$ along a circle around the $z-$axis.

In spherical polar coordinates, $h_r = 1$ again, and $h_\theta = r$. A change $\mathrm{d}\phi$ in ϕ corresponds to moving a distance $r\sin\theta\, \mathrm{d}\phi$ (because $r\sin\theta$ is the radius of the particular latitudinal circle around the $z-$axis), so $h_\phi = r\sin\theta$.

One reason that orthogonal coordinates are so useful is that in any orthogonal coordinate system (u_1, u_2, u_3), small displacements along u_1 and u_2 define small rectangles, while small displacements along u_1, u_2, u_3 define small cuboids. In other words, $h_1 h_2\, du_1 du_2$ is an area element normal to $\mathbf{e}_3$ on a surface of constant u_3, and $h_1 h_2 h_3\, du_1 du_2 du_3$ is a volume element.

5.6 Vector fields and vector algebra in curvilinear coordinates

Scalar fields can of course be expressed in (orthogonal) curvilinear coordinates: they are simply written as functions $f(u_1, u_2, u_3)$ or for brevity $f(u_i)$.

As you will know from Linear Algebra, vectors can be expressed using any basis of the vector space concerned. The same is true, at each point, of vector fields. Up to now we have always chosen $\mathbf{i}, \mathbf{j}, \mathbf{k}$ as our basis vectors: however, when using curvilinear coordinates we will normally use the orthogonal unit vectors along the coordinate lines as our basis vectors, and write

$$\mathbf{F} = F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3.$$

For clarity, we can use the coordinate names instead of 1,2,3 as subscripts for the three components. Thus we may write

$$\begin{aligned} \mathbf{F} &= F_x\mathbf{i} + F_y\mathbf{j} + F_z\mathbf{k} \\ &= F_\rho\mathbf{e}_\rho + F_\phi\mathbf{e}_\phi + F_z\mathbf{e}_z \\ &= F_r\mathbf{e}_r + F_\theta\mathbf{e}_\theta + F_\phi\mathbf{e}_\phi . \end{aligned}$$

to express the same vector in Cartesian, cylindrical polar and spherical polar coordinates (of course $\mathbf{e}_x = \mathbf{i}$ and so on in Cartesians). Note that the **same** vector $\mathbf{F}$ will have **different** components depending on our choice of basis vectors: suppose we are given an $\mathbf{F}$ with defined F_x, F_y, F_z above, but we want to find F_r, F_θ, F_ϕ, then we need to use the matrix as in Eq. 5.3 to express $\mathbf{i}, \mathbf{j}, \mathbf{k}$ in terms of the $\mathbf{e}$'s, multiply out and collect into one term in each $\mathbf{e}$. (This effectively turns into a matrix multiplication).

In any orthogonal coordinate system, the scalar (dot) and vector (cross) products work just as in Cartesian coordinates:

$$\mathbf{w}.\mathbf{v} = w_1v_1 + w_2v_2 + w_3v_3 \tag{5.5}$$

and

$$\mathbf{w} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ w_1 & w_3 & w_3 \\ v_1 & v_2 & v_3 \end{vmatrix} , \tag{5.6}$$

but note this only works if the vectors are defined **at the same point** , such as a dot product $\mathbf{F} \cdot d\mathbf{r}$ or $\mathbf{F} \cdot d\mathbf{S}$ in a line or surface integral. We **cannot** use these for two position vectors at widely separate points, because the $\mathbf{e}$'s vary with position.

Vector differentiation is more complicated, because the unit vectors are no longer constant: when we differentiated a vector in Cartesians

$$\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$$

we just differentiated the components (F_1, F_2, F_3) because the unit vectors are constant; but in general coordinates the $\mathbf{e}$'s depend on position, so we have to use the product rule and differentiate the $\mathbf{e}$ vectors as well as the components F_i.

Differentiation of these vectors with respect to a variable other than position (like the derivatives in Section 3.1) is straightforward. For example if position $\mathbf{r}$ depends on time, and is given in cylindrical polars so $\mathbf{r} = \rho\mathbf{e}_\rho + z\mathbf{e}_z$, we just use the product rule to get the time derivative

$$\dot{\mathbf{r}} = \dot{\rho}\mathbf{e}_\rho + \rho\dot{\mathbf{e}}_\rho + \dot{z}\mathbf{e}_z + z\dot{\mathbf{e}}_z \ .$$

(where the over-dots are shorthand for time derivative, as is common). Then since $\mathbf{e}_\rho = \cos\phi\mathbf{i} + \sin\phi\mathbf{j}$ from (5.2),

$$\dot{\mathbf{e}}_\rho = \dot{\phi}(-\sin\phi\mathbf{i} + \cos\phi\mathbf{j}) = \dot{\phi}\mathbf{e}_\phi \ .$$

Similarly $\dot{\mathbf{e}}_z = \mathbf{0}$. Substituting into the previous result, we get

$$\dot{\mathbf{r}} = \dot{\rho}\mathbf{e}_\rho + \rho\dot{\phi}\mathbf{e}_\phi + \dot{z}\mathbf{e}_z$$

for a velocity in cylindrical polar coordinates.

When differentiating scalar or vector fields with respect to position, the key operations are always grad of a scalar, and div and curl of a vector field (this is because these are the only combinations that behave "sensibly" after rotations). In the next sections, we will show how to calculate the grad, div and curl operators in general orthogonal coordinates; then we apply those general formulae to the most common cases of cylindrical polars and spherical polars.

5.7 The Gradient Operator in curvilinear coordinates

To calculate the gradient of a scalar field $V(u_1, u_2, u_3)$ in orthogonal curvilinear coordinates (u_1, u_2, u_3), we go back to the definition

$$\mathrm{d}V = \nabla V \cdot \mathrm{d}\mathbf{r} \qquad . \qquad (*)$$

for the change dV caused by an infinitesimal position change $d\mathbf{r}$.

(**Note:** here dV is the infinitesimal change in scalar field V resulting from a small change $d\mathbf{r}$; it is not a volume element.)

We define $\nabla V \equiv (\nabla V)_1\mathbf{e}_1 + (\nabla V)_2\mathbf{e}_2 + (\nabla V)_3\mathbf{e}_3$, and we want to find the three components $(\nabla V)_1$ etc.

From the definitions of the unit vectors previously, we have $\mathrm{d}\mathbf{r} = \mathbf{e}_1 h_1 \mathrm{d}u_1 + \mathbf{e}_2 h_2 \mathrm{d}u_2 + \mathbf{e}_3 h_3 \mathrm{d}u_3$, so the right-hand side of $(*)$ becomes

$$\begin{aligned}((\nabla V)_1\mathbf{e}_1 + (\nabla V)_2\mathbf{e}_2 + (\nabla V)_3\mathbf{e}_3)\cdot \qquad & (\mathbf{e}_1 h_1 \mathrm{d}u_1 + \mathbf{e}_2 h_2 \mathrm{d}u_2 + \mathbf{e}_3 h_3 \mathrm{d}u_3) \\ & = (\nabla V)_1 h_1 \mathrm{d}u_1 + (\nabla V)_2 h_2 \mathrm{d}u_2 + (\nabla V)_3 h_3 \mathrm{d}u_3\end{aligned}$$

using the orthogonality of the **e**'s.

Now turning to the left-hand side of of $(*)$, using Taylor's theorem (in 3 dimensions), and discarding terms of second and higher derivatives, we get

$$dV = \frac{\partial V}{\partial u_1}\mathrm{d}u_1 + \frac{\partial V}{\partial u_2}\mathrm{d}u_2 + \frac{\partial V}{\partial u_3}\mathrm{d}u_3$$

These two expressions above must be equal for **any** arbitrary changes $\mathrm{d}u_1$, $\mathrm{d}u_2$ and $\mathrm{d}u_3$. Hence we must have

$$(\nabla V)_1 h_1 = \frac{\partial V}{\partial u_1}\ ; \quad (\nabla V)_2 h_2 = \frac{\partial V}{\partial u_2}\ ; \quad (\nabla V)_3 h_3 = \frac{\partial V}{\partial u_3} \quad .$$

Dividing by the h's and substituting back into the original definition, in orthogonal curvilinear coordinates we have

$$\nabla V \quad = \quad \frac{1}{h_1}\frac{\partial V}{\partial u_1}\mathbf{e}_1 + \frac{1}{h_2}\frac{\partial V}{\partial u_2}\mathbf{e}_2 + \frac{1}{h_3}\frac{\partial V}{\partial u_3}\mathbf{e}_3 \quad . \qquad (5.7)$$

Clearly in Cartesian coordinates, we have $u_1 = x$, $\mathbf{e}_1 = \mathbf{i}$ etc and all three h's are 1, so this simplifies to the well-known formula from Chapter 1.

For a geometrical explanation, the $1/h_i$ terms take care of the arc-length effects, i.e. how far $\mathbf{r}$ moves for a small change in each coordinate. So the 1-component of ∇V represents the change dV per small **distance**

ds in the direction $\mathbf{e}_1$; but, moving a distance ds in direction $\mathbf{e}_1$ requires a change $\delta u_1 = ds/h_1$ in coordinate u_1; therefore the $1/h_i$ terms appear in grad V above.

Example 5.6. What is ∇V in spherical polar coordinates ? Evaluate ∇V where $V = r\sin\theta\cos\phi$.

In spherical polars, $(u_1,u_2,u_3) = (r,\theta,\phi)$ and $h_1 = 1$, $h_2 = r$, $h_3 = r\sin\theta$. Putting those into 5.7 we have

$$\nabla V = \frac{\partial V}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial V}{\partial \theta}\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial V}{\partial \phi}\mathbf{e}_\phi \quad .$$

For the given V, $\partial V/\partial r = \sin\theta\cos\phi$, $\partial V/\partial\theta = r\cos\theta\cos\phi$ and $\partial V/\partial\phi = -r\sin\theta\sin\phi$. Hence, using the result above,

$$\nabla V = \sin\theta\cos\phi\,\mathbf{e}_r + \cos\theta\cos\phi\,\mathbf{e}_\theta - \sin\phi\,\mathbf{e}_\phi .$$

(In this case we can observe that $V = x$ and $\nabla V = \mathbf{i}$, using the matrix from Eq. 5.3, so this example is a lot easier in Cartesians; however, many problems involving circular or spherical symmetry do get easier in polar coordinates).

Exercise 5.1. What is ∇V in cylindrical polar coordinates (ρ,ϕ,z) ? □

Exercise 5.2. Let (r,θ,ϕ) be spherical polar coordinates. Evaluate ∇f where

$$\text{(a) } f = \phi\,; \quad \text{(b) } f = \theta\,; \quad \text{(c) } f = (r^n\sin m\theta)\,.$$

□

5.8 The Divergence Operator in curvilinear coordinates

Next we want to compute $\nabla\cdot\mathbf{F}$ in orthogonal curvilinear coordinates. Although we could directly calculate the divergence in any coordinates, using the Cartesian definition, the matrix relating basis unit vectors, and the chain rule, the results can be found with less effort from the Divergence Theorem. The Divergence Theorem is true in all coordinates (since it equates scalars, whose value must be independent of the coordinates). Thus

$$\int_V \nabla\cdot\mathbf{F}\,\mathrm{d}V = \int_{\mathscr{S}} \mathbf{F}.\mathrm{d}\mathbf{S}\,,$$

where $\mathscr{S}$ is the closed surface enclosing volume V.

Now, we apply this to an infinitesimal "cuboid" with one corner at (u_1,u_2,u_3) and edges corresponding to changes δu_1, δu_2, δu_3 in each coordinate; so this has eight corners at $(u_1,u_2,u_3),(u_1+\delta u_1,u_2,u_3),\ldots(u_1+\delta u_1,u_2+\delta u_2,u_3+\delta u_3)$. From before, the volume of the cuboid is $\delta V = (h_1\delta u_1)(h_2\delta u_2)(h_3\delta u_3)$. For a sufficiently small volume, we can approximate $\nabla\cdot\mathbf{F}$ as constant across δV, so the left-hand side becomes

$$(\nabla\cdot\mathbf{F})\delta V = (\nabla\cdot\mathbf{F})(h_1h_2h_3\delta u_1\delta u_2\delta u_3)\,.$$

Next we consider the right-hand side of the Divergence Theorem: we need to take the surface integral over the six faces of our cuboid, and add results. First consider the integral of $\mathbf{F}.\mathbf{n}$ over the face of the cuboid where the first coordinate has value $u_1+\delta u_1$. This face is a rectangle with unit normal $+\mathbf{e}_1$ and area $(h_2\delta u_2)(h_3\delta u_3)$, so the surface integral is approximately

$$(h_2h_3\delta u_2\delta u_3F_1)_{u_1+\delta u_1}\,,$$

where the subscript shows it is evaluated at $u_1+\delta u_1$. On the opposite face at u_1 we have unit normal $-\mathbf{e}_1$ (pointing outwards i.e. away from the first face), so the surface integral gives us

$$-(h_2h_3\delta u_2\delta u_3F_1)_{u_1}\ .$$

Repeating the above for the other four faces we get symmetrical results; finally summing the six terms and then taking the limit as $\delta V\to 0$, we obtain

$$\begin{aligned}\nabla\cdot\mathbf{F} \quad &= \lim_{\delta u_1,\delta u_2,\delta u_3\to 0}\frac{1}{\delta V}\Big[\quad (h_2\delta u_2h_3\delta u_3F_1)_{u_1+\delta u_1}-(h_2\delta u_2h_3\delta u_3F_1)_{u_1}\\ &\qquad +(h_3\delta u_3h_1\delta u_1F_2)_{u_2+\delta u_2}-(h_3\delta u_3h_1\delta u_1F_2)_{u_2}\\ &\qquad +(h_1\delta u_1h_2\delta u_2F_3)_{u_3+\delta u_3}-(h_1\delta u_1h_2\delta u_2F_3)_{u_3}\Big]\ .\end{aligned}$$

Though each pair of brackets looks the same, this is not zero because the h's and F's are different on opposite faces of the cuboid; the first two terms give us δu_1 times the partial derivative $\partial/\partial u_1$ of the bracket, and so on for the next pairs, so this gets us the result

$$\nabla\cdot\mathbf{F}=\frac{1}{h_1h_2h_3}\left[\frac{\partial(h_2h_3F_1)}{\partial u_1}+\frac{\partial(h_3h_1F_2)}{\partial u_2}+\frac{\partial(h_1h_2F_3)}{\partial u_3}\right]. \tag{5.8}$$

Note: In this last step, we have taken some δu's outside the brackets and cancelled them with the ones in δV, but we must leave the h's **inside** the differentiation since the h's generally vary with position. This comes about because our "cuboid" may be slightly "tapering", so the areas of opposite faces are not exactly equal; and differentiating the h_i's takes care of that.

Example 5.7. What is $\nabla\cdot\mathbf{F}$ in cylindrical polar coordinates, where $\mathbf{F}=F_\rho\mathbf{e}_\rho+F_\phi\mathbf{e}_\phi+F_z\mathbf{e}_z$?

In cylindrical polars, $(u_1,u_2,u_3)=(\rho,\phi,z)$ and $h_1=1$, $h_2=\rho$, $h_3=1$. Hence

$$\nabla\cdot\mathbf{F}=\frac{1}{\rho}\left[\frac{\partial(\rho F_\rho)}{\partial\rho}+\frac{\partial F_\phi}{\partial\phi}+\frac{\partial(\rho F_z)}{\partial z}\right].$$

Note that we can apply the product rule, and since $\partial\rho/\partial z=0$, $\partial\rho/\partial\rho=1$ we get

$$\nabla\cdot\mathbf{F}=\frac{1}{\rho}F_\rho+\frac{\partial F_\rho}{\partial\rho}+\frac{1}{\rho}\frac{\partial F_\phi}{\partial\phi}+\frac{\partial F_z}{\partial z}.$$

Note: It is important to note that an F_ρ term has appeared here, which is not a derivative of F. This has appeared because the coordinate lines for ρ have a "built in divergence", they all radiate outwards from the z-axis, so a field with constant F_ρ has a positive divergence term due to this.

As a further example we can note that in cylindrical polars, $\mathbf{r}=\rho\mathbf{e}_\rho+0\mathbf{e}_\phi+z\mathbf{e}_z$. Plugging in components $(\rho,0,z)$ to the above, we get

$$\nabla\cdot\mathbf{r}=1+1+0+1=3$$

which agrees with the result in Cartesians, as it must.

(If we had just taken $\partial\rho/\partial\rho+\partial z/\partial z$ we would have got $\nabla\cdot\mathbf{r}=2$; clearly wrong) .

Example 5.8. What is $\nabla\cdot\mathbf{F}$ in spherical polar coordinates, where $\mathbf{F}=F_r\mathbf{e}_r+F_\theta\mathbf{e}_\theta+F_\phi\mathbf{e}_\phi$?

In spherical polars, $(u_1,u_2,u_3)=(r,\theta,\phi)$ and $h_1=1$, $h_2=r$, $h_3=r\sin\theta$. Hence

$$\nabla\cdot\mathbf{F}=\frac{1}{r^2\sin\theta}\left[\frac{\partial(r^2\sin\theta F_r)}{\partial r}+\frac{\partial(r\sin\theta F_\theta)}{\partial\theta}+\frac{\partial(rF_\phi)}{\partial\phi}\right].$$

5.9 The Curl Operator in curvilinear coordinates

Finally we want curl: as before we have curvilinear coordinates (u_1, u_2, u_3), and a vector field $\mathbf{F} = F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3$; we want to calculate

$$\nabla \times \mathbf{F} \equiv (\nabla \times \mathbf{F})_1\mathbf{e}_1 + (\nabla \times \mathbf{F})_2\mathbf{e}_2 + (\nabla \times \mathbf{F})_3\mathbf{e}_3 \ ,$$

so we want the 1,2,3 components of the above.

In analogy with the previous section, we use Stokes's theorem to provide a coordinate-independent definition of $\nabla \times \mathbf{F}$:

$$\int_{\mathscr{S}} (\nabla \times \mathbf{F}) \cdot \mathrm{d}\mathbf{S} = \int_{\mathscr{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r},$$

where $\mathscr{S}$ is a surface spanning the closed curve $\mathscr{C}$.

To calculate the 1-component $(\nabla \times \mathbf{F})_1$, consider a planar curve around a small "rectangle" on a surface of constant u_1, with sides given by small changes δu_2 and δu_3. From previous results, the vector area of this rectangle $d\mathbf{S} = h_2\delta u_2 h_3 \delta u_3\, \mathbf{e}_1$; now taking $(\nabla \times \mathbf{F}) \cdot d\mathbf{S}$, the 2 and 3 components of $\nabla \times \mathbf{F}$ disappear so the LHS of Stokes's theorem is approximately

$$(\nabla \times \mathbf{F})_1 h_2 \delta u_2 h_3 \delta u_3 \quad .$$

Now looking at the RHS of Stokes's theorem, the line integral around the edge of the same rectangle is given by adding the line integrals along the four sides: this is approximately

$$(h_2\delta u_2 F_2)_{u_3} + (h_3\delta u_3 F_3)_{u_2+\delta u_2} - (h_2\delta u_2 F_2)_{u_3+\delta u_3} - (h_3\delta u_3 F_3)_{u_2} \ ,$$

where the subscripts denote that the term is evaluated at that value, and two minus signs appear because opposite sides are traversed in opposite directions around the closed rectangle. Equating the last two expressions, and taking the limit as δu_2, $\delta u_3 \to 0$, we have

$$\begin{aligned}(\nabla \times \mathbf{F})_1 &= \frac{1}{h_2 h_3} \lim_{\delta u_2, \delta u_3 \to 0} \left[\frac{(h_3F_3)_{u_2+\delta u_2} - (h_3F_3)_{u_2}}{\delta u_2} - \frac{(h_2F_2)_{u_3+\delta u_3} - (h_2F_2)_{u_3}}{\delta u_3} \right] \\ &= \frac{1}{h_2 h_3} \left(\frac{\partial (h_3F_3)}{\partial u_2} - \frac{\partial (h_2F_2)}{\partial u_3} \right).\end{aligned}$$

This is just the 1-component of $\nabla \times \mathbf{F}$. To get the 2- and 3- components, we just repeat all the above for two more small rectangles in surfaces of constant u_2, u_3 respectively; this looks the same but cycling the 1/2/3's, and we get

$$(\nabla \times \mathbf{F})_2 = \frac{1}{h_3 h_1} \left(\frac{\partial (h_1F_1)}{\partial u_3} - \frac{\partial (h_3F_3)}{\partial u_1} \right),$$

$$(\nabla \times \mathbf{F})_3 = \frac{1}{h_1 h_2} \left(\frac{\partial (h_2F_2)}{\partial u_1} - \frac{\partial (h_1F_1)}{\partial u_2} \right).$$

These results can be written in a compact (and more memorable) form as a determinant:

$$\nabla \times \mathbf{F} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1\mathbf{e}_1 & h_2\mathbf{e}_2 & h_3\mathbf{e}_3 \\ \partial/\partial u_1 & \partial/\partial u_2 & \partial/\partial u_3 \\ h_1F_1 & h_2F_2 & h_3F_3 \end{vmatrix}. \tag{5.9}$$

Once again, in Cartesian coordinates this simplifies to the well-known expression from Chapter 3.4.

Example 5.9. What is $\nabla \times \mathbf{F}$ in spherical polar coordinates?

In spherical polar coordinates (r,θ,ϕ) we have $h_1 = 1$, $h_2 = r$, $h_3 = r\sin\theta$. Hence, using the determinant form:

$$\nabla\times\mathbf{F} = \frac{1}{r^2\sin\theta}\begin{vmatrix} \mathbf{e}_r & r\mathbf{e}_\theta & r\sin\theta\mathbf{e}_\phi \\ \partial/\partial r & \partial/\partial\theta & \partial/\partial\phi \\ F_r & rF_\theta & r\sin\theta F_\phi \end{vmatrix}.$$

or in expanded form

$$\nabla\times\mathbf{F} = \frac{1}{r^2\sin\theta}\left[\frac{\partial(r\sin\theta F_\phi)}{\partial\theta} - \frac{\partial(rF_\theta)}{\partial\phi}\right]\mathbf{e}_r + \frac{1}{r\sin\theta}\left[\frac{\partial F_r}{\partial\phi} - \frac{\partial(r\sin\theta F_\phi)}{\partial r}\right]\mathbf{e}_\theta + \frac{1}{r}\left[\frac{\partial(rF_\theta)}{\partial r} - \frac{\partial F_r}{\partial\theta}\right]\mathbf{e}_\phi\ .$$

Note that since r is independent of θ and ϕ, etc., we can for instance take the r outside the differentiations in the $\mathbf{e}_r$ component and cancel it with an r in the denominator. Remember the answer is a curl so it's a vector field. Do not add all the components together, forgetting the vectors $\mathbf{e}_r$ etc (this is a common error).

Note: the full expression above looks quite daunting. However in many problems this may simplify considerably using symmetry: for example, if a given problem is symmetrical around the z-axis, then we will have $F_\phi = 0$ and $\partial F_r/\partial\phi = 0$ and $\partial F_\theta/\partial\phi = 0$, so four of the six derivatives will vanish.

Exercise 5.3. Show by expanding it that the determinant definition is equivalent to the full expressions for the individual components given above. □

Exercise 5.4. What is $\nabla\times\mathbf{F}$ in cylindrical polar coordinates?

Note that if ρ and z have dimensions of length and ϕ is dimensionless (because it's an angle), then all the terms in the expression for $\nabla\times\mathbf{F}$ should have the same dimensions, namely the dimensions of $\mathbf{F}$ divided by length. This is a simple check that you should make. □

Exercise 5.5. Use spherical polar coordinates to evaluate the divergence and curl of $\mathbf{r}/r^3$. [Hint: don't forget that in spherical polar coordinates, the position vector $\mathbf{r}$ is equal to $r\mathbf{e}_r$.] □

Exercise 5.6. State Stokes's theorem, and verify it for the hemispherical surface $r = 1$, $z \geq 0$, with the vector field $\mathbf{A}(\mathbf{r}) = (y,-x,z)$. □

Exercise 5.7. The vector field $\mathbf{B}(\rho) = (0,\rho^{-1},0)$ in cylindrical polar coordinates (ρ,ϕ,z). Evaluate $\nabla\times\mathbf{B}$. Evaluate the line integral $\int_{\mathscr{C}}\mathbf{B}.\mathrm{d}\mathbf{r}$, where $\mathscr{C}$ is the unit circle $z = 0$, $\rho = 1$, $0 \leq \phi \leq 2\pi$. Does Stokes's theorem apply? □

Note: To conclude this chapter, we will note that many applied maths or Physics problems involve an expression like $\nabla^2 V$, where V is a scalar field and ∇^2 is the Laplacian operator, in cylindrical or spherical polar coordinates. We can get the expressions for $\nabla^2 V$ in polar coordinates using firstly the definition Eq. 3.10 (recall this was $\nabla^2 V \equiv \mathrm{div}(\mathrm{grad}\,V)$) , and then using Eq. 5.7 for grad V, then taking div of that with Eq. 5.8.

The results are available in most textbooks; you will not be expected to memorise those, but you might be given them in an exam question and asked to calculate something, so it's worth taking a look especially if you are taking applied maths courses later.

Appendix

Other ways of doing Example 5.5 are as follows

The second method is to divide the volume removed into two parts: (i) a cylinder with radius $\sin\theta_1$ and height $\cos\theta_1$, and (ii) a 'top-slice'. Volume (i), the cylinder, is easy: $2\pi\sin^2\theta_1\cos\theta_1$. To get volume (ii) we integrate over ρ and then z

$$4\pi\int_{\cos\theta_1}^{1} dz\int_0^{\sqrt{1-z^2}}\rho\,d\rho = 2\pi\int_{\cos\theta_1}^{1}(1-z^2)\,dz = \frac{2\pi}{3}(2+\cos^3\theta_1-3\cos\theta_1).$$

The sum of volumes (i) and (ii) is $\frac{4\pi}{3}(1-\cos^3\theta_1)$ as expected.

A third way also divides the volume removed into two parts: (i) an 'ice-cream cone' or cone with a spherical top, and (ii) a cylinder minus cone. The volume (i) is

$$4\pi\int_0^{\theta_1}\sin\theta\,d\theta\int_0^1 r^2\,dr = \frac{4\pi}{3}(1-\cos\theta_1).$$

Volume (ii), a cylinder with cone removed, is a bit harder:

$$4\pi\int_0^{\cos\theta_1} dz\int_{z\tan\theta_1}^{\sin\theta_1}\rho\,d\rho = 2\pi\int_0^{\cos\theta_1}(\sin^2\theta_1 - z^2\tan^2\theta_1)\,dz = \frac{4\pi}{3}\sin^2\theta_1\cos\theta_1$$

(which notice is $\frac{2}{3}$ of the volume of the cylinder). Again the sum of the volumes integrated is $\frac{4\pi}{3}(1-\cos^3\theta_1)$.

Finally, a fourth possibility is to integrate for the volume remaining after coring, which is

$$4\pi\int_0^{\cos\theta_1} dz\int_{\sin\theta_1}^{\sqrt{1-z^2}}\rho\,d\rho = 2\pi\int_0^{\cos\theta_1}(1-z^2-\sin^2\theta_1)\,dz = \frac{4\pi}{3}\cos^3\theta_1.$$

SUMMARY OF ORTHOGONAL CURVILINEAR COORDINATES

In orthogonal curvilinear coordinates (u_1, u_2, u_3), with corresponding unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and arc-length parameters h_1, h_2, h_3, the gradient of a scalar field V is given by

$$\nabla V = \frac{1}{h_1}\frac{\partial V}{\partial u_1}\mathbf{e}_1 + \frac{1}{h_2}\frac{\partial V}{\partial u_2}\mathbf{e}_2 + \frac{1}{h_3}\frac{\partial V}{\partial u_3}\mathbf{e}_3 \qquad ;$$

the divergence of a vector field $\mathbf{F} = F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3$ is given by

$$\nabla\cdot\mathbf{F} = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial u_1}(h_2h_3F_1) + \frac{\partial}{\partial u_2}(h_3h_1F_2) + \frac{\partial}{\partial u_3}(h_1h_2F_3)\right] \qquad ;$$

and the curl of the same vector field is given by

$$\nabla\times\mathbf{F} = \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\mathbf{e}_1 & h_2\mathbf{e}_2 & h_3\mathbf{e}_3 \\ \partial/\partial u_1 & \partial/\partial u_2 & \partial/\partial u_3 \\ h_1F_1 & h_2F_2 & h_3F_3 \end{vmatrix} \qquad .$$

Cartesian coordinates:
$(u_1, u_2, u_3) \equiv (x, y, z)$; arc-length parameters $h_1 = 1$, $h_2 = 1$, $h_3 = 1$.

Cylindrical polar coordinates:
$(u_1, u_2, u_3) \equiv (\rho, \phi, z)$; arc-length parameters $h_1 = 1$, $h_2 = \rho$, $h_3 = 1$.

Spherical polar coordinates:
$(u_1, u_2, u_3) \equiv (r, \theta, \phi)$; arc-length parameters $h_1 = 1$, $h_2 = r$, $h_3 = r\sin\theta$.

Chapter 6

Fourier series

Last updated: 01 Dec 2010.
Syllabus section:
6. Fourier series: full, half and arbitrary range series. Parseval's Theorem.

Fourier series provide a way to do various calculations with, and to analyse the behaviour of, functions which are **periodic**: this means that they repeat the same values in a regular pattern, or are defined in a finite range. Specifically, a function which is **periodic** with period L will obey an equation

$$f(x+L) = f(x) \qquad \text{for all } x$$

and to start with, we will assume $L = 2\pi$ also for convenience. We already know that $\cos nx$ and $\sin nx$ for any integer n have period 2π. (So, of course, do the other trigonometric functions such as $\tan x$, but these have the disadvantage of becoming unbounded at certain values, e.g. $\tan x$ is unbounded at $x = \pi/2$).

The basic principle of Fourier series is to express our periodic function $f(x)$ as an infinite sum of sine and cosine functions,

$$f(x) = \sum_{0}^{\infty} (a_n \cos nx + b_n \sin nx)$$

for a periodic and piecewise differentiable $f(x)$ (in fact, for any function defined on a range of length 2π). We will slightly modify this way of writing the series soon.

Such a series splits f into pieces of different "frequency": geometrically, each of the $\sin nx$ and $cosnx$ terms has exactly n positive and negative"wiggles" over the range $0 \leq x \leq 2\pi$, and the a_n, b_n are constants telling us how much f varies at each different frequency.

This technique (and its generalisation to Fourier transforms) has a large number of practical applications, including: resolution of sound waves into their different frequencies, e.g. in MP3 players; telecommunications and Wi-Fi; computer graphics and image processing; astronomy and optics; climate variation; water waves; periodic behaviour of financial measures, etc.

6.1 Full range Fourier series

As above, the idea is that we have a given function $f(x)$ defined for a range of values of x of length 2π, say $-\pi \le x \le \pi$; now we approximate this function as an infinite sum of trigonometric functions, as

$$f(x) \approx S(x) \equiv \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx. \tag{6.1}$$

where a_n, b_n are an infinite series of constants to be determined. The right-hand side of this, $S(x)$ for short, is called the **Fourier series** for $f(x)$, and the set of coefficients a_n, b_n are called the **Fourier coefficients**. Here the $\frac{1}{2}a_0$ is really a $\cos 0x = 1$ constant term, and the $\frac{1}{2}$ is put in for convenience as we see below. (There is no point in including a b_0 term since $\sin 0x = 0$).

Clearly, to make progress we have to actually calculate the a_n, b_n; this looks very hard since we there are infinitely many of them, but is actually straightforward using the **orthogonality** properties of $\sin mx, \cos nx$: the key results we need are, for any two non-negative integers m and n,

$$\begin{aligned}
\int_{-\pi}^{\pi} \cos mx \sin nx \, dx &= 0 \\
\int_{-\pi}^{\pi} \cos mx \cos nx \, dx &= \begin{cases} 0 & \text{if } m \neq n \\ \pi & \text{if } m = n \neq 0 \\ 2\pi & \text{if } m = n = 0 \end{cases} \\
\int_{-\pi}^{\pi} \sin mx \sin nx \, dx &= \begin{cases} 0 & \text{if } m \neq n \\ \pi & \text{if } m = n \neq 0 \\ 0 & \text{if } m = n = 0 \end{cases}
\end{aligned} \tag{6.2}$$

All of the above are simple to prove using the trigonometric identities from Chapter 1, e.g. $\cos A \cos B = \frac{1}{2}[\cos(A+B) + \cos(A-B)]$ and similar. Using these, we can find the Fourier coefficients given $f(x)$: suppose we multiply Eq. 6.1 by $\cos mx$ for some *fixed* integer m, then integrate from $-\pi$ to π, then we have

$$\int_{-\pi}^{\pi} f(x) \cos mx \, dx = \int_{-\pi}^{\pi} \left[\tfrac{1}{2}a_0 \cos mx + \sum_{1}^{\infty} a_n \cos nx \cos mx + \sum_{1}^{\infty} b_n \sin nx \cos mx \right] dx$$

Assuming the sums converge, we can swap the integral sign and the summations above, giving

$$\int_{-\pi}^{\pi} f(x) \cos mx \, dx = \tfrac{1}{2}a_0 \left[\int_{-\pi}^{\pi} \cos mx \, dx \right] + \sum_{n=1}^{\infty} a_n \left[\int_{-\pi}^{\pi} \cos nx \cos mx \, dx \right] + \sum_{n=1}^{\infty} b_n \left[\int_{-\pi}^{\pi} \sin nx \cos mx \, dx \right] \tag{6.3}$$

Now suppose $m > 0$, and look at the integrals in square-brackets above: the first one is zero. From Eq. 6.2, the integrals in the middle term are all zero, except for exactly one case when $n = m$ when the integral is π. The integrals in the right-hand term are all zeros. Therefore, the RHS of the above is simply one non-zero term $= a_m \pi$; so rearranging we get

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx \, \mathrm{d}x \, .$$

Likewise, if instead we multiplied Eq. 6.1 by $\sin mx$ and integrated, we get

$$\int_{-\pi}^{\pi} f(x) \sin mx \, dx = \tfrac{1}{2}a_0 \left[\int_{-\pi}^{\pi} \sin mx \, dx \right] + \sum_{n=1}^{\infty} a_n \left[\int_{-\pi}^{\pi} \cos nx \sin mx \, dx \right] + \sum_{n=1}^{\infty} b_n \left[\int_{-\pi}^{\pi} \sin nx \sin mx \, dx \right]$$

Again all the square-brackets on the RHS are zero, except for one case in the rightmost bracket with $n = m$ which gives π; so the RHS is $b_m \pi$ and we rearrange to

$$b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx \, \mathrm{d}x \, .$$

Finally, we need the special case of $m = 0$: going back to Eq. 6.3 the LHS contains $\cos mx = \cos 0x = 1$; now the a_0 term on the RHS is the only one which gives a non-zero integral, because both the infinite sums have $n \geq 1 \neq m$ and all the integrals are zero. Then the RHS above becomes $\frac{1}{2}a_0(2\pi)$, so the above equation for a_m is still correct for $m = 0$; note that the funny-looking $\frac{1}{2}$ in the original definition Eq. 6.1 was put in to make that work. (Some books may not have the $\frac{1}{2}$ in Eq. 6.1, but then we need to add a $\frac{1}{2}$ in the equation defining a_0 instead). Remember $\sin 0x = 0$ so there is no b_0 term to deal with.

The equations above were derived by choosing one **fixed** integer m and showing that all terms with $n \neq m$ disappeared: however the argument is correct for any value of m, so the above equations give **all** the coefficients a_m, b_m. (The choice of letter m above is arbitrary, but it had to be different to the n which runs from 0 to ∞). Finally, since m is a dummy label in the above and n has now disappeared, we can change the letter m back to n and we get

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x \qquad (n \geq 0) \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,\mathrm{d}x \qquad (n \geq 1) \end{aligned} \tag{6.4}$$

Therefore, **to find the Fourier series** $S(x)$ for a given $f(x)$, we simply have to evaluate the definite integrals Eq. 6.4 (using a suitable method such as integration by parts) to get a_n, b_n for all n; then substitute those coefficients back into Eq. 6.1.

Next we take an example of actually evaluating the a_n, b_n for a given $f(x)$.

Example 6.1. Find the Fourier series for

$$f(x) = \begin{cases} 0 & \text{if } -\pi < x < 0 \\ x & \text{if } 0 < x < \pi. \end{cases}$$

Using the formulae above,

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x = \frac{1}{\pi}\int_{0}^{\pi} x\cos nx\,\mathrm{d}x \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,\mathrm{d}x = \frac{1}{\pi}\int_{0}^{\pi} x\sin nx\,\mathrm{d}x \end{aligned}$$

(the lower limits become 0 because we were given $f(x) = 0$ in $[-\pi, 0]$, so that range contributes zero to the integrals). Evaluating the above, using integration by parts, we find that:

$$\begin{aligned} a_n &= \frac{1}{\pi}\left(\left[\frac{x\sin nx}{n}\right]_0^{\pi} - \int_0^{\pi}\frac{\sin nx}{n}\,\mathrm{d}x\right) = \frac{1}{\pi}\left[\frac{\cos nx}{n^2}\right]_0^{\pi} \\ &= \frac{1}{\pi n^2}(\cos n\pi - 1) \\ &= \frac{1}{\pi n^2}((-1)^n - 1) \end{aligned}$$

and this gives $a_n = -2/\pi n^2$ when n is odd, or $a_n = 0$ for even $n > 0$.

Note that for $n = 0$ the procedure above contains $0/0$ so is ill-defined: as is common, we need to treat $n = 0$ as a special case, with $\cos 0x = 1$:

$$a_0 = \frac{1}{\pi}\int_0^{\pi} x\,1\,\mathrm{d}x = \frac{1}{\pi}\left[\frac{x^2}{2}\right]_0^{\pi} = \frac{\pi}{2}.$$

Finally we need the b_n's, which are

$$\begin{aligned} b_n &= \frac{1}{\pi}\left(\left[\frac{-x\cos nx}{n}\right]_0^\pi + \int_0^\pi \frac{\cos nx}{n}\,\mathrm{d}x\right) = \frac{1}{\pi}\left(-\frac{\pi\cos n\pi}{n} + \left[\frac{\sin nx}{n^2}\right]_0^\pi\right) \\ &= \frac{1}{\pi n}(-\pi\cos n\pi) + 0 = \frac{-(-1)^n}{n} \\ &= \frac{(-1)^{n+1}}{n} \end{aligned}$$

(and there is no b_0 term, so this gives b_n for all positive n).

Putting all these a_n, b_n back into the general form Eq. 6.1, the Fourier series we are asked for is

$$S(x) = \frac{\pi}{4} - \sum_{k=0}^{\infty} \frac{2}{\pi(2k+1)^2}\cos(2k+1)x + \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin nx\,.$$

where we have dealt with the odd/even n for a_n by replacing n with $2k+1$ which must be odd, and summing over $k = 0$ to ∞.

Although this general method always works (as long as we can evaluate the integrals), we do not need to do it for functions we can put into the required form by other means, as in the next example.

Example 6.2. Find the Fourier series for $\sin^4 x$.

Here we use the double angle formula: $\sin^4 x = \frac{1}{4}(1-\cos 2x)^2 = \frac{1}{4}(1 - 2\cos 2x + \cos^2 2x) = \frac{1}{4}(1 - 2\cos 2x + \frac{1}{2}[1+\cos 4x])$
so $\sin^4 x = \frac{3}{8} - \frac{1}{2}\cos 2x + \frac{1}{8}\cos 4x$.
This already looks like a special case of Eq. 6.1, so we just write $a_0 = \frac{3}{4}$ (remembering the half), $a_2 = -\frac{1}{2}$, $a_4 = \frac{1}{8}$; and all other a_n and all b_n are zero.

(**Note:** We could evaluate the integrals and get the same answer, but we don't need to do that here since we can see the result by inspection).

We note that the series $S(x)$ is periodic, i.e. if we take the same series for any x, rather than staying in the range $-\pi \le x \le \pi$, $S(x)$ will obey $S(x+2\pi) = S(x)$. So this can also be used for functions defined on a range longer than 2π if those functions are periodic with period 2π. Another way to look at this is that if we know the function on the range $[-\pi, \pi]$ we can define it for all x by insisting that it be periodic; graphically, this is equivalent to just "copying" the function infinitely many times for intervals 2π, like wallpaper.

We note that the range of x could equally well be $[\alpha, \alpha+2\pi]$ for any α, since all the quantities involved are periodic so this will give integrals over exactly the same range of values of f. Note $\alpha = 0$ is often used, so the range of x becomes $[0, 2\pi]$.

Exercise 6.1. Find the Fourier series of $f(x)$ defined by $f(x) = 0$ in $-\pi < x < 0$ and $f(x) = \cos x$ in $0 \le x < \pi$.
The answer should be

$$\tfrac{1}{2}\cos x + \sum_{k=1}^{\infty}\frac{4k}{\pi(4k^2-1)}\sin 2kx\,.$$

□

Going back to example 6.1, and evaluating both sides at $x = \pi/2$: we need to remember that the cosine of an odd multiple of $\pi/2$ is zero, the sine of an even multiple of $\pi/2$ is zero, and the sine of $(2k+1)\pi/2$ is

$(-1)^{k-1}$, from chapter 1. So we get

$$\begin{aligned} \frac{\pi}{2} &= \frac{\pi}{4} + \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \\ \Rightarrow \frac{\pi}{4} &= 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots \end{aligned}$$

A number of results of this sort, giving sums of numerical series, can be obtained by direct evaluation of equation (6.1) at some particular x. The only tricky point in using this is to guess which x to evaluate: usually one of π, $\pi/2$ or $\pi/4$ is what is needed, to make the sine and cosine functions give simple results such as 0 or $(-1)^n$ etc.

Warning: so far, we have not actually proved that the infinite sum $S(x)$ on the right-hand side of 6.1 actually converges, or has limit $f(x)$. Strictly, what we have shown is that IF there exists an infinite sum $S(x)$ which does converge to $f(x)$ over $-\pi \leq x \leq \pi$, then the coefficients must be given by Eq. 6.4.

We discuss the question of convergence and the limit in the next section.

6.2 Completeness and convergence of Fourier series

We now give answers to two questions: can every function with period 2π be written this way, and does the series $S(x)$ in 6.1 with coefficients 6.4 always converge at all x ? These ideas are referred to as completeness and convergence. To specify more fully, consider the sum of the first N terms with x fixed: this sum definitely exists since all the a_n, b_n are bounded if $f(x)$ is bounded, and we get a sum of a finite N bounded terms). Then let $N \to \infty$: if the limit exists, then $S(x)$ is said to converge at x. Completeness amounts to asking if this limit $S(x)$ equals the value of the original function $f(x)$. The proof of the relevant properties is not part of this course, but the result is. As usual, the conditions in it are like small print in contracts – ignorable most of the time, but important when things go wrong.

Theorem 6.1 *(Fourier's theorem or Dirichlet's theorem) If $f(x)$ is periodic with period 2π for all x, and $f(x)$ is piecewise smooth in $(-\pi, \pi)$, then the Fourier series $S(x)$ with coefficients a_n and b_n (defined as above) converges to $\frac{1}{2}(f(x+) + f(x-))$ at every point.*

Here "piecewise smooth" means sufficiently differentiable at all except isolated points, and $f(x+)$ means the limit of $f(x+\delta)$ as δ (positive) tends to zero, which is called the upper limit or right limit of $f(x)$ at x. Similarly $f(x-)$ is the limit of $f(x-\delta)$ as δ tends to zero, called the lower limit or left limit). At any x where $f(x)$ is continuous, we have $f(x+) = f(x-) = f(x)$, so $S(x) = \frac{1}{2}[f(x) + f(x)] = f(x)$ so the Fourier series does converge to exactly $f(x)$. At points where $f(x)$ has a discontinuity, $f(x+)$ and $f(x-)$ are not equal, and then $S(x) = \frac{1}{2}(f(x+) + f(x-))$ gives the average value of $f(x)$ on either side of the discontinuity: but this may not be the value of $f(x)$ itself at the point.

Typically, we will find that as $n \to \infty$, the coefficients a_n and b_n tend to zero like $1/n$ or faster.

Example 6.3. Taking the function and series of Example 6.1, Fourier's theorem tells us that at $x = \pi$ the series converges to $\frac{1}{2}(f(\pi+) + f(\pi-)) = \frac{1}{2}(0 + \pi) = \frac{1}{2}\pi$, using $f(\pi+) = f((-\pi)+)$ by periodicity. The series then gives

$$\frac{\pi}{2} = \frac{\pi}{4} + \sum_{k=0}^{\infty} \frac{2}{\pi(2k+1)^2} \, ,$$

since $\sin n\pi = 0$ and $\cos(2k+1)\pi = -1$. Subtracting $\pi/4$ we have

$$\begin{aligned} \frac{\pi}{4} &= \sum_{k=1}^{\infty} \frac{2}{\pi(2k+1)^2} = \frac{2}{\pi}(1+\frac{1}{3^2}+\frac{1}{5^2}\ldots), \quad \text{therefore} \\ \frac{\pi^2}{8} &= 1+\frac{1}{3^2}+\frac{1}{5^2}+\ldots \quad . \end{aligned}$$

As a nice corollary of the above, we can get the infinite sum for all integers (not just odd ones) as follows: define

$$T \equiv 1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots \qquad ,$$

then dividing by 4 gives

$$\frac{1}{4}T = \frac{1}{2^2}+\frac{1}{4^2}+\frac{1}{6^2}+\ldots$$

so subtracting,

$$\frac{3}{4}T = 1+\frac{1}{3^2}+\frac{1}{5^2}+\ldots$$

which is the series above. Therefore

$$T = \frac{4}{3}\frac{\pi^2}{8} = \frac{\pi^2}{6}$$

Note: There is a strange detail. Fourier's theorem tells us what happens in the limit of the infinite series. But if we take any finite number of terms we obviously cannot match a discontinuity exactly, since the finite series must give a continuous function. It turns out that any finite sum overshoots the function on either side of the discontinuity: this curious effect is called Gibbs's phenomenon— adding more terms does not reduce the overshoot, it just moves the overshoot closer to the discontinuity. (In the limit of the infinite sum, the overshoot gets "infinitesimally close" to the discontinuity, so for any x a finite distance from the discontinuity, this does not matter).

Example 6.4. The square wave.

Consider the "square wave" function defined by

$$f(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 \text{ if } x > 0 \end{cases} \tag{6.5}$$

in the domain $[-\pi, \pi]$ and periodic with period 2π. This gives

$$a_0 = 1 \qquad a_{n>0} = 0 \qquad b_n = \frac{1-\cos n\pi}{n\pi}$$

so b_n is 0 for even n or $2/(n\pi)$ for odd n. Therefore,

$$f(x) = \tfrac{1}{2} + 2\sum_{n \text{ odd}} \frac{\sin nx}{n\pi} \, . \tag{6.6}$$

Figure 6.1 shows the square wave and its approximations by its Fourier series (up to $n = 1$ and $n = 5$). Several things are noticeable:

(i) even a square wave, which looks very unlike sines and cosines, can be approximated by them, to any desired accuracy;

(ii) although we only considered the domain $[-\pi, \pi]$ the Fourier series automatically extends the domain to all real x by generating a periodic answer;

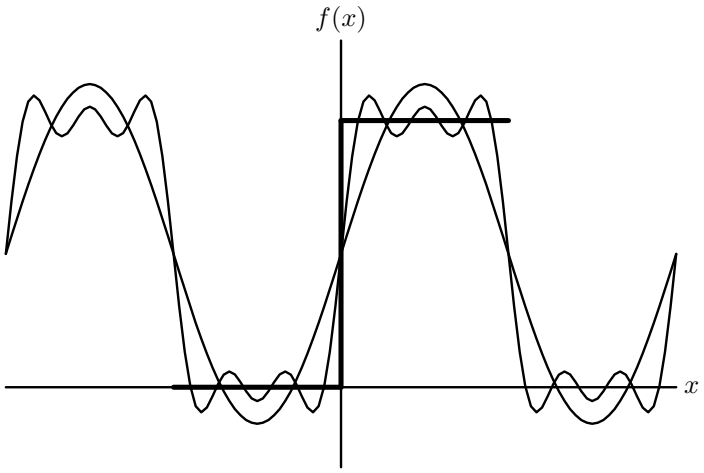

Figure 6.1: Square wave (as in equation (6.5) but with the vertical direction stretched for better visibility) and Fourier partial sums: two terms and four terms.

(iii) at discontinuities, the Fourier series gives the mean value of $f(x)$ on either side of the discontinuity.

(iv) close to discontinuities the Fourier series overshoots.

Another result telling us in what sense we have a good approximation is **Parseval's theorem:**

Theorem 6.2 *(Parseval's Theorem) If $f(x)$ has a Fourier series defined as in Section 6.1, then*

$$\int_{-\pi}^{\pi} f(x)^2 \,\mathrm{d}x = \tfrac{1}{2}\pi a_0^2 + \pi \sum_{n=1}^{\infty} (a_n^2 + b_n^2).$$

For a formal proof one has to deal with convergence of the infinite sum, but if we assume convergence we can write

$$f(x)^2 = \left(\tfrac{1}{2}a_0 + \sum_1^\infty a_n \cos nx + \sum_1^\infty b_n \sin nx\right)\left(\tfrac{1}{2}a_0 + \sum_1^\infty a_m \cos mx + \sum_1^\infty b_m \sin mx\right)$$

then we can expand this out into a double sum

$$\begin{aligned} f(x)^2 \quad = \quad & \frac{1}{4}a_0^2 + \tfrac{1}{2}a_0\left(\sum_1^\infty a_m \cos mx + \sum_1^\infty b_m \sin mx\right) + \tfrac{1}{2}a_0\left(\sum_1^\infty a_n \cos nx + \sum_1^\infty b_n \sin nx\right) \\ & + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}(a_n a_m \cos nx \cos mx + a_n b_m \cos nx \sin mx + b_n a_m \sin nx \cos mx + b_n b_m \sin nx \sin mx) \end{aligned}$$

(Note: in the above, *n* and *m* can be any letters, but we have to use two **different** letters since we're summing over both of them independently).

Now as before we integrate the above from $x = -\pi$ to π, and again we swap the sum and integral signs: the first term is a constant giving integral $(1/4)a_0^2 2\pi$; the next two terms contain only single sin's and cos's which all integrate to zero. Then in the double sum, we look up results from Eq.6.2 again, and all the terms with $m \neq n$ integrate to zero: so we can turn the double summation into a single summation with $m = n$ (think

of summing over an infinite chessboard where all off-diagonal squares contain zeros). Then, the $\sin mx \cos nx$ terms also integrate to zero: finally the $\cos mx \cos nx$ terms and $\sin mx \sin nx$ terms (with $m = n$) integrate to π, so the overall result is

$$\int_{-\pi}^{\pi} f(x)^2 \, dx = \frac{1}{4} a_0^2 (2\pi) + 0 + 0 + \sum_{n=1}^{\infty} (a_n^2 \pi + 0 + 0 + b_n^2 \pi) \quad ;$$

this is Parseval's theorem as above.

In a very similar way, one can show that for two functions $f(x)$ and $g(x)$, with $f(x)$ having Fourier coefficients a_n, b_n and $g(x)$ having coefficients A_n, B_n, we obtain

$$\int_{-\pi}^{\pi} f(x) g(x) \, \mathrm{d}x = \tfrac{1}{2} \pi a_0 A_0 + \pi \sum_{n=1}^{\infty} (a_n A_n + b_n B_n) \; .$$

Example 6.5. Go back to the Fourier series for the square wave, Eq. 6.5 above. Putting this into both sides of Parseval's theorem, we have

$$\begin{aligned} \int_0^{\pi} 1 \, \mathrm{d}x &= \frac{\pi}{2} + \frac{4\pi}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{(2k+1)^2} \\ \pi &= \frac{\pi}{2} + \frac{4}{\pi} (1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots) \end{aligned}$$

On rearranging we get

$$\frac{\pi^2}{8} = \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots$$

which we had already derived in another way in Example 6.3.

Parseval's theorem is important in practical applications, for example telling us numerically "how good" is an approximation to $f(x)$ given by taking only a finite number of terms in the Fourier series (as we have to do in real-world evaluation on a computer). We proceed as follows: define $S_N(x)$ to be the sum up to and including $n = N$ of the Fourier series for $f(x)$, then $S(x)$ is the infinite sum (the limit of $S_N(x)$ as N tends to infinity). If we define $E_N(x) = f(x) - S_N(x)$, this is the "residual error" if we keep only the first N terms of the series.

It is easy to see that the Fourier series for $E_N(x)$ has coefficients zero for $1 \le n \le N$, and a_n, b_n for $n > N$, so applying Parseval's theorem to $E_N(x)$,

$$\int_{-\pi}^{\pi} (E_N(x))^2 \, dx = \pi \sum_{n=N+1}^{\infty} (a_n^2 + b_n^2)$$

If we divide the above equation by the range 2π, the left-hand side becomes the mean value of E_N^2 over the range, which is the "mean square error" in our approximation $S_N(x)$. So, if the right-hand side is small, i.e. the sum of $a_n^2 + b_n^2$ is converging rapidly to its limit, we know that $S_N(x)$ is a good approximation of our original function $f(x)$.

6.3 Odd and even functions; Half range Fourier series

We recall the definitions of an "even" and "odd" function:
$f(x)$ is even $\Leftrightarrow f(x) = f(-x)$ for all x.
$f(x)$ is odd $\Leftrightarrow f(x) = -f(-x)$ for all x.

Any function $f(x)$ can always be written as

$$f(x) = \tfrac{1}{2}[f(x) + f(-x)] + \tfrac{1}{2}[f(x) - f(-x)] \ ,$$

in which the first bracket on the right is an even function and the second bracket is an odd function, by construction.

Since $\sin kx$ is odd and $\cos kx$ is even, we might suspect that for even functions $f(x)$ only cosine terms appear in the Fourier series (all $b_n = 0$), while similarly for odd functions only sine terms appear and all $a_n = 0$. This is correct, and we can easily check this, e.g.

$$\begin{aligned} \pi a_n &= \int_{-\pi}^{\pi} f(x)\cos nx \, \mathrm{d}x \\ &= \int_{-\pi}^{0} f(x)\cos nx \, \mathrm{d}x + \int_{0}^{\pi} f(x)\cos nx \, \mathrm{d}x \\ &= \int_{u=\pi}^{0} f(-u)\cos(-nu)\,(-1)\mathrm{d}u + \int_{0}^{\pi} f(x)\cos nx \mathrm{d}x \end{aligned}$$

where we have substituted $u = -x$ in the first half, so its range becomes π to 0. Now this is

$$\begin{aligned} &= -\int_{0}^{\pi} f(-u)\cos nu\,(-1)\mathrm{d}u + \int_{0}^{\pi} f(x)\cos nx \mathrm{d}x \\ &= \int_{0}^{\pi} (f(-x) + f(x))\cos nx \mathrm{d}x \end{aligned}$$

where we have replaced u by $+x$ since it's a dummy variable. The above is clearly 0 if $f(x)$ is an odd function.

Similarly

$$\pi b_n = \int_{0}^{\pi} (f(x) - f(-x))\sin nx \mathrm{d}x.$$

To summarise the above, if $f(x)$ is an even function, we have

$$a_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\cos nx \, \mathrm{d}x \ , \quad b_n = 0 \text{ for all } n \quad .$$

(where by symmetry we can halve the range of integration from 0 to π, and multiply by 2). And if $f(x)$ is an odd function , all $a_n = 0$, and

$$b_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\sin nx \, \mathrm{d}x \ .$$

We can use this property to make a Fourier series for a **half range** using **only** sine or only cosine terms, as follows. Suppose we are given a function $\phi(x)$ defined on $[0, \pi]$ (a "half range"), then we can define two new functions on the range $[-\pi, \pi]$: we construct an even function $f(x)$ such that $f(x) = \phi(x)$ in $(0, \pi)$ and $f(x) = \phi(-x)$ if $-\pi < x < 0$. Likewise we define an odd function $g(x)$ such that $g(x) = \phi(x)$ for $0 \le x < \pi$, and $g(x) = -\phi(-x)$ if $-\pi < x < 0$.

Note that **both** $f(x)$ and $g(x)$ are equal to $\phi(x)$ on the range $(0, \pi)$, but they have opposite signs on the range $(-\pi, 0)$. (Note also that $h(x) = \frac{1}{2}(f(x) + g(x))$ is equal to $\phi(x)$ on $(0, \pi)$ and zero on $(-\pi, 0)$).

Inserting these $f(x)$ and $g(x)$ into Eq. 6.1, our even function $f(x)$ gives a Fourier series with

$$a_n = \frac{2}{\pi}\int_{0}^{\pi} \phi(x)\cos nx \, \mathrm{d}x \ , \qquad b_n = 0,$$

and the odd function $g(x)$ gives a Fourier series with

$$a_n = 0\ , \qquad b_n = \frac{2}{\pi}\int_0^\pi \phi(x)\sin nx\,\mathrm{d}x\ .$$

These are called respectively the **half-range cosine series** and **half-range sine series** for $\phi(x)$; both of those series are equal to $\phi(x)$ on the range $(0,\pi)$, but they have opposite signs on the range $(-\pi,0)$.

(Also it is clear that if you take the average of the above two series, you get the series for $h(x)$ above, which is equal to $\phi(x)$ on $(0,\pi)$ and zero on $(-\pi,0)$).

Example 6.6. $f(x)$ is such that $f(x)=f(x+2\pi)$ and $f(x)=-f(-x)$, and on $0\le x\le\pi$, $f(x)=x(\pi-x)$. Find its Fourier series, and prove that

$$1-\frac{1}{3^3}+\frac{1}{5^3}+\ldots=\frac{\pi^3}{32}$$

The given $f(x)$ has period 2π and is odd, so we know the series contains only sine terms, and

$$\begin{aligned}
b_n &= \frac{2}{\pi}\int_0^\pi x(\pi-x)\sin nx\,\mathrm{d}x \\
&= \frac{2}{\pi}\left\{\left[-x(\pi-x)\frac{\cos nx}{n}\right]_0^\pi + \int_0^\pi (\pi-2x)\frac{\cos nx}{n}\,\mathrm{d}x\right\} \\
&= \frac{2}{\pi}\left\{\left[(\pi-2x)\frac{\sin nx}{n^2}\right]_0^\pi + 2\int_0^\pi \frac{\sin nx}{n^2}\,\mathrm{d}x\right\} \\
&= \frac{4}{\pi}\left[-\frac{\cos nx}{n^3}\right]_0^\pi \\
&= \begin{cases} 0 & \text{for } n=2k, \\ \dfrac{8}{\pi(2k+1)^3} & \text{for } n=2k+1. \end{cases}
\end{aligned}$$

Thus

$$f(x)=\frac{8}{\pi}\sum_{k=0}^{\infty}\frac{\sin(2k+1)x}{(2k+1)^3}. \tag{6.7}$$

To get the series requested, we try evaluating (6.7) at some x such that $\sin(2k+1)x=(-1)^k$. This occurs at $x=\pi/2$. Evaluating both sides there gives

$$f(\pi/2)=\frac{\pi^2}{4}=\frac{8}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)^3}$$

which on rearranging gives the required result.

6.4 Arbitrary range Fourier series

Here we extend the Fourier series to the case when the range of our function is not $-\pi\le x\le\pi$. If we have $f(x)$ defined in a range $-L\le x\le L$, instead of $-\pi<x\le\pi$, then we can define a new variable $y\equiv\pi x/L$ (a rescaled version of x), so that $-\pi\le y\le\pi$ and write f as a Fourier series in y.

$$\begin{aligned}
f(x) &= \tfrac{1}{2}a_0+\sum_{n=1}^{\infty}(a_n\cos ny+b_n\sin ny) \\
&= \tfrac{1}{2}a_0+\sum_{n=1}^{\infty}(a_n\cos(\frac{n\pi x}{L})+b_n\sin(\frac{n\pi x}{L})).
\end{aligned}$$

where

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{y=-\pi}^{y=\pi} f(\frac{Ly}{\pi})\cos\frac{n\pi x}{L}\mathrm{d}(\frac{\pi x}{L}), \\ &= \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}\mathrm{d}x. \end{aligned}$$

and similarly

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}\mathrm{d}x.$$

Here we have just "rescaled": observe that as x goes from $-L$ to L, the quantity $n\pi x/L$ goes from -πn to $+\pi n$ so there are again an integer n "wiggles" in each cos/sin term.

For functions which are a simple stretch/squash of another function whose Fourier series we have already worked out, we can rescale variables.

Example 6.7. Find the Fourier series for the function $g(x)$ of period $2c$ such that

$$g(x) = \begin{cases} 0 & \text{if } -c < x < 0 \\ x & \text{if } 0 < x < c. \end{cases}$$

Using the result of example 6.1, replacing x by y, we have

$$\begin{aligned} f(y) &= \frac{\pi}{4} - \sum_{k=1}^{\infty}\frac{2}{\pi(2k+1)^2}\cos(2k+1)y + \sum_{1}^{\infty}\frac{(-1)^{n+1}}{n}\sin ny \qquad -\pi < y \le \pi \\ \Rightarrow f(\frac{\pi x}{c}) &= \frac{\pi}{4} - \sum_{k=1}^{\infty}\frac{2}{\pi(2k+1)^2}\cos\frac{(2k+1)\pi x}{c} + \sum_{1}^{\infty}\frac{(-1)^{n+1}}{n}\sin\frac{n\pi x}{c} \qquad -c \le x \le c \end{aligned}$$

But we have $f(\pi x/c) = 0$ for $-c < x < 0$, or $\pi x/c$ for $0 < x \le c$, so $f(\pi x/c) = (\pi/c)g(x)$ for all $-c < x \le c$. So we just multiply the series above by c/π, and get

$$\Rightarrow g(x) = \frac{c}{4} - \sum_{k=1}^{\infty}\frac{2c}{\pi^2(2k+1)^2}\cos\frac{(2k+1)\pi x}{c} + \sum_{1}^{\infty}\frac{(-1)^{n+1}c}{\pi n}\sin\frac{n\pi x}{c}$$

Appendix

This section will not be lectured and is not for examination

The following shows the kind of application Fourier himself had in mind and gives an example of some methods in partial differential equations which we will meet in another context in the next chapter.

Example 6.8. In the propagation of heat in a solid in one dimension, the temperature θ obeys the equation

$$k\frac{\partial^2\theta}{\partial x^2} = \frac{\partial\theta}{\partial t}.$$

This is the simplest case of the **diffusion equation**.

We introduce here a new idea which will run through the rest of the course. This is **separation of variables**: we can see that if we look for a solution in the form $X(x)T(t)$ we will find

$$kT\frac{\mathrm{d}^2X}{\mathrm{d}x^2} = X\frac{\mathrm{d}T}{\mathrm{d}t} \Rightarrow \frac{k}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2} = \frac{1}{T}\frac{\mathrm{d}T}{\mathrm{d}t}.$$

Here the left side depends only on x and the right side only on t: hence the two sides must both equal the same constant (only a constant can depend only on x, and only on t, at the same time). We then have two equations

$$k\frac{\mathrm{d}^2X}{X\,\mathrm{d}x^2} = \lambda, \qquad \lambda = \frac{\mathrm{d}T}{T\,\mathrm{d}t},$$

to solve, where λ is our unknown constant. When we have solved these, we multiply the answers together to solve the original equation. In general we assume (and indeed usually we can prove) that the full solution is a (possibly infinite) sum of solutions of the separable type.

For Fourier's problem we proceed as follows:
At the earth's surface, the temperature θ is assumed to vary periodically over the year (for simplicity) so it has a Fourier series in time t with period 1 year. We define x to be the depth into the earth. Then at the surface $x = 0$ we can write

$$\theta = \tfrac{1}{2}a_0 + \sum_{n=1} (a_n \cos\frac{2n\pi t}{T} + b_n \sin\frac{2n\pi t}{T})$$

with $T = 365/2$ days.

Now at other x we let a_n and b_n depend on x and put these into the differential equation: this means we are writing the whole solution as a sum of separable solutions in which the t dependence gives a Fourier series (with different coefficients at each x). Plugging this into the original equation and equating coefficients in the Fourier series we get

$$k\frac{\partial^2 a_n}{\partial x^2}\cos\frac{2n\pi t}{T} = \frac{2n\pi}{T}b_n\cos\frac{2n\pi t}{T}.$$

$$k\frac{\partial^2 b_n}{\partial x^2}\sin\frac{2n\pi t}{T} = -\frac{2n\pi}{T}a_n\sin\frac{2n\pi t}{T}.$$

These can be written as a single complex equation

$$\frac{\partial^2(b_n + \mathrm{i}a_n)}{\partial x^2} = \frac{2n\pi\mathrm{i}}{T}(b_n + \mathrm{i}a_n).$$

This equation is easy to solve as it is a linear equation with constant coefficients. [For those who have done the Differential Equations course, the auxiliary equation has roots

$$\pm\sqrt{\frac{n\pi}{kT}}(1+i)$$

and that gives the solutions. We need the solution with a negative real part (temperature variation decreases as we go into the earth).] The solution is

$$b_n + \mathrm{i}a_n = c\exp\left(-\sqrt{\frac{n\pi}{kT}}(1+i)x\right),$$

for some constant c. This means we have a solution which varies sinusoidally with time, but the amplitude of variation decreases by a factor e in a distance $\sqrt{kT/n\pi}$. Some realistic figures are $k = 2.10^{-3}$ cm^2/s, $T = 365.24.3600/2$ secs, giving $1/\lambda \equiv \sqrt{kT/\pi} = 177$ cm for annual variation and roughly $1/19$ of this for daily variation. The amplitude of the annual variation halves in a distance x such that $\lambda x = \ln 2$, about 123 cm. So in 5 metres the variation of temperature reduces by a factor 1/16 (it also turns out that at that depth the variation is out of phase with the surface, i.e. coolest in mid-summer).

6.5 Fourier Transforms

This section is not examinable, but is included since it may be useful for later courses.

To conclude this chapter, it is worth a quick look at the extension of Fourier series to Fourier Transforms. The principle remains the same, i.e. expressing a general function as a sum of trigonometric functions of different frequency.

There are two main steps to get from Fourier series to Fourier transforms: firstly, we introduce complex numbers and use Euler's formula

$$e^{inx} = \cos nx + i \sin nx \qquad .$$

Then we change the definition of the Fourier series to

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

and the coefficients c_n become

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx$$

What we have done here is just make the coefficients c_n complex, extended the infinite sum to negative integers n, and changed the prefactor from $1/\pi$ to $1/(2\pi)$ to compensate for doubling the number of terms in the sum. (The $n = 0$ case does not have positive and negative terms so the half in Eq. 6.1 gets absorbed in the above). In this case we can easily see, taking real and imaginary parts of the above, that $c_n = \frac{1}{2}(a_n - ib_n)$ where the a_n, b_n are the same as previous sections; assuming $f(x)$ is real-valued, then it is clear from the definition that $c_{-n} = \frac{1}{2}(a_n + ib_n) = \bar{c_n}$, the complex conjugate.

This has not really done anything very new, it just turns two real formulae for a_n, b_n into one complex formula for c_n. The real parts of the c_n's are the cosine terms and the imaginary parts give the sine terms; if we extract the two terms for $+n$ and $-n$ in the series for $f(x)$ we have

$$\begin{aligned} c_n e^{inx} + c_{-n} e^{-inx} &= \tfrac{1}{2}(a_n - ib_n)(\cos nx + i \sin nx) + \tfrac{1}{2}(a_n + ib_n)(\cos nx - i \sin nx) && (6.8) \\ &= (a_n \cos nx + b_n \sin nx) \quad , && (6.9) \end{aligned}$$

so the imaginary parts cancel, and this agrees with what we had before.

This also allows us to extend the formula to complex-valued $f(x)$, in which case the terms $c_n + \bar{c}_{-n}$ are no longer real, and their imaginary parts give the complex part of $f(x)$.

To extend to Fourier transforms, we generalise the above to the arbitrary-range series, i.e. let $f(x)$ be periodic with period L, i.e.

$$\begin{aligned} f(x) &= \sum_{-\infty}^{\infty} c_n e^{-2\pi inx/L} \\ c_n &= \frac{1}{L} \int_{-L/2}^{L/2} f(x) e^{2\pi inx/L}\, dx \end{aligned}$$

Now if we write $\delta = 2\pi/L$ and $\omega_n = n\delta$ this becomes

$$\begin{aligned} f(x) &= \sum_{-\infty}^{\infty} c_n e^{i\omega x} \\ c_n &= \frac{\delta}{2\pi} \int_{-L/2}^{L/2} f(x) e^{i\omega x}\, dx \end{aligned}$$

and if we let the range L tend to infinity, let $d_n = c_n/\delta$, let δ tend to zero, we can convert the infinite discrete series of coefficients d_n into a continuous function $\mathscr{F}(\omega)$, and (skipping some details) we arrive at

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} \mathscr{F}(\omega) e^{i\omega x}\, d\omega \\ \mathscr{F}(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-i\omega x}\, dx \qquad . \end{aligned} \tag{6.10}$$

Here $\mathscr{F}(\omega)$ is called the **Fourier transform** of $f(x)$, with ω called the (angular) frequency , which is the continuous version of the n we had before.

Note: there are several possible "arbitrary choices" of where to put the 2π's and minus signs in the above definitions; some books put a factor $1/\sqrt{2\pi}$ before both integrals, which makes them symmetrical. Other authors leave a 2π inside the exponential term, in which case ω is usually changed to a different letter e.g. $\nu = \omega/2\pi$. As long as this is done consistently, it doesn't matter, but there must be factors of 2π somewhere in the definitions.

Chapter 7

Laplace's Equation

Last update: 13 Dec 2010.
Syllabus section;
7. Laplace's equation. Uniqueness under suitable boundary conditions. Separation of variables. Two-dimensional solutions in Cartesian and polar coordinates. Axisymmetric spherical harmonic solutions.

7.1 The Laplace and Poisson equations

Let $\Phi(\mathbf{r})$ be a scalar field in three dimensions, as in previous chapters. Laplace's equation is simply

$$\nabla^2\Phi = 0 \tag{7.1}$$

where, as we met in Chapter 3.6, $\nabla^2\Phi \equiv \nabla\cdot(\nabla\Phi) \equiv \text{div}(\text{grad}\,\Phi)$; here ∇^2 is called the **Laplacian operator**, or just the Laplacian.

Remember from before, if Φ is a scalar field, its gradient $\nabla\Phi$ is a vector field, and then taking div of that gives us another scalar field: so Laplace's equation is a scalar equation.

In Cartesian x,y,z coordinates, things are simple: we recall the definitions from Chapter 3,

$$\nabla\Phi = \frac{\partial\Phi}{\partial x}\mathbf{i} + \frac{\partial\Phi}{\partial y}\mathbf{j} + \frac{\partial\Phi}{\partial z}\mathbf{k}$$

and

$$\nabla\cdot\mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$$

Putting $\mathbf{F} = \nabla\Phi$ above, so $F_1 = \partial\Phi/\partial x$ etc, Laplace's equation in Cartesians is

$$\nabla^2\Phi \equiv \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = 0 \tag{7.2}$$

Note that if we are using other coordinates (e.g. cylindrical polars or spherical polars) we must use results for grad and div in those coordinates from Chapter 5, so it will look different; we look at those later.

Laplace's equation often occurs as follows: suppose we have a conservative vector field $\mathbf{F}$, so that $\mathbf{F} = \nabla\Phi$ for some scalar field Φ as in Chapter 4.7; then if $\nabla\cdot\mathbf{F} = 0$ this gives Laplace's equation $\nabla^2\Phi = 0$.

Aside: Laplace's equation is the simplest and most basic example of one of the three types of second-order linear partial differential equations (PDEs), known as the *"elliptic"* type. Laplace's equation is a linear homogeneous equation.

A generalisation of Laplace's equation is **Poisson's equation** which is

$$\nabla^2\Phi = f(\mathbf{r}) \quad ,$$

where $f(\mathbf{r})$ is a given scalar field. Laplace's equation is clearly a special case of Poisson's where $f(\mathbf{r}) = 0$ at all points in the volume of interest.

The basic examples of the other types of PDE are the **wave equation**

$$\frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = \nabla^2 f \ ,$$

where c is constant (usually the speed of sound or light) and t is time;
and the **heat equation** or **diffusion equation**

$$\frac{\partial f}{\partial t} = \kappa\nabla^2 f$$

where f is temperature in a solid, and κ is a constant. (We met the heat equation with a single spatial variable in Example 6.8 on Fourier series).

In maths, the wave equation is an example of a *"hyperbolic"* PDE and and the heat equation is a *"parabolic"* PDE. These names are potentially confusing since the solutions have nothing to do with ellipses, parabolas, or hyperbolas, but this is just a "shorthand" because the powers and signs in the equations look rather similar to the equations for ellipsoids, paraboloids and hyperboloids from Chapter 1.

Laplace's and Poisson's equations are very important, both because of their occurrence in many physics applications, and because they are the basic examples of elliptic PDEs. We are now going to spend the rest of this chapter considering some solutions of Laplace's equation in 2 dimensions.

We can see directly that there are some simple solutions of Laplace's equation, e.g.

$$\begin{aligned}
\Phi &= c \qquad \text{constant} \\
\Phi &= x \\
\Phi &= y \\
\Phi &= xy \\
\Phi &= x^2 - y^2 \\
&\text{etc}
\end{aligned}$$

These clearly are solutions, by direct evaluation of $\nabla^2\Phi$ from Eq. 7.2. There are in fact an infinite number of general solutions to Laplace's equation, which are known as **harmonic functions**.

We easily see that ∇^2 is a linear operator: that is

$$\nabla^2(\lambda\Phi_1 + \mu\Phi_2) = \lambda\nabla^2\Phi_1 + \mu\nabla^2\Phi_2 \quad .$$

for any two scalar fields Φ_1, Φ_2 and any two constants λ, μ (independent of position), since both grad and div have this property. Hence if Φ_1 and Φ_2 are both solutions of Laplace's equation, so is $\lambda\Phi_1 + \mu\Phi_2$. Also, if Ψ is a solution of Poisson's equation and Φ is a solution of Laplace's equation, $\Psi + \Phi$ is also a solution of Poisson's equation, for the same $f(\mathbf{r})$.

Aside: Considering some gravitational and electromagnetic examples of conservative fields, and using the Divergence Theorem

$$\int_V \nabla\cdot\mathbf{F}\,\mathrm{d}V = \int_S \mathbf{F}.\mathrm{d}\mathbf{S}$$

we see that if $\nabla \cdot \mathbf{F} = 0$ everywhere there are no sources inside the volume, which for gravity means that there is no mass there, and for electric field means that there is no (net) charge. Hence, Laplace's equation describes the gravitational potential in regions of space where there is no matter, and the electric potential in regions where there are no charges.

If instead there is a net charge density ρ, the electric field $\mathbf{E}$ satisfies

$$\nabla \cdot \mathbf{E} = \frac{1}{\varepsilon_0}\rho(\mathbf{r})$$

where ε_0 is a constant of nature. (This is one of the four Maxwell's equations). Combining this with $\mathbf{E} = -\nabla\Phi$ gives

$$\nabla^2\Phi = -\frac{1}{\varepsilon_0}\rho(\mathbf{r}).$$

That is an example of Poisson's equation as we met above. Laplace's equation is of course a special case of Poisson's equation, in which the function on the right-hand side is zero throughout the volume of interest.

7.2 Uniqueness of Solutions to Poisson's (and Laplace's) Equation

Here, we will prove that under suitable boundary conditions the solution of Poisson's (or Laplace's) equation is **unique**. We shall then investigate what the solutions actually are in some simple cases, in each of Cartesian, cylindrical and spherical polar coordinates.

As is common in differential equations, there are many general solutions (in fact an infinite family), so to find the solution in a specific case we need to be given some **boundary conditions**. Recall for a 1-D ordinary differential equation we often need a function value at one or two ends of a line; but here since Laplace's equation works in 3 dimensions, usually we need the value of $\Phi(\mathbf{r})$ to be given at all points on a closed surface S, and we solve Laplace's equation to find Φ in the volume **inside** S. (Occasionally we solve over the infinite volume outside S, with another boundary condition for Φ at infinity).

Theorem 7.1 *Suppose that $\nabla^2 U = f(\mathbf{r})$ throughout some closed volume V, $f(\mathbf{r})$ being some specified function of $\mathbf{r}$, and that the value of U is specified at every point on the surface S bounding volume V. Then, if a solution $U(\mathbf{r})$ exists to this problem, it is unique.*

Proof:
Before proceeding, we need to recall Eq. 3.6, which was

$$\nabla \cdot (U\mathbf{F}) = U\nabla \cdot \mathbf{F} + (\nabla U) \cdot \mathbf{F} \qquad .$$

Choosing $\mathbf{F} = \nabla U$ in the above, we get the identity

$$\nabla \cdot (U\nabla U) = U\nabla^2 U + (\nabla U) \cdot (\nabla U) \qquad (*)$$

which we use below.

Now to prove the uniqueness theorem, suppose that U_1 and U_2 are two scalar fields which *both* solve the given problem. Define $W \equiv U_1 - U_2$ to be the difference of our two solutions.

Then, we know that $\nabla^2 W = 0$ inside volume V (by linearity), and W = 0 at all points on the surface S, since both U_1 and U_2 match the given boundary condition.

Now we consider the volume integral

$$\begin{aligned}
\int_V |\nabla W|^2 \,\mathrm{d}V &= \int_V (\nabla W).(\nabla W)\,\mathrm{d}V \\
&= \int_V \nabla\cdot(W\nabla W) - W\nabla^2 W \;\mathrm{d}V \qquad \text{using } (*) \text{ above} \\
&= \int_V \nabla\cdot(W\nabla W)\; dV \; - 0 \qquad \text{since } \nabla^2 W = 0 \text{ everywhere in } V \\
&= \int_S (W\nabla W)\cdot \mathrm{d}\mathbf{S} \qquad (\text{by the Divergence Theorem}) \\
&= 0 \qquad \text{because } W = 0 \text{ on } S
\end{aligned}$$

Now, the integrand on the LHS is a squared quantity, therefore is always non-negative, and its integral is zero. This can only happen if $\nabla W = \mathbf{0}$ throughout V (otherwise, if ∇W was non-zero anywhere in V, the whole integral on the LHS will be positive because there cannot be any negative bits in the integrand to cancel the positive part, i.e. a contradiction).

Now $\nabla W = \mathbf{0}$ throughout V means W is a constant throughout V. But $W = 0$ on the boundary of V, therefore $W = 0$ throughout V. Hence $U_1 = U_2$ throughout V, so the solution is **unique**. Q.E.D.

Note that we have actually proved uniqueness for Poisson's equation, and Laplace's is a special case of that.

[**Aside:** It is fairly clear that the final step in the displayed calculation above also works if, instead of $W = 0$ on the boundary, $\nabla W \cdot \mathbf{n} = 0$ where $\mathbf{n}$ is the normal to the surface S. This corresponds to being given a boundary condition for $\nabla U \cdot \mathbf{n}$ on the boundary, instead of the value of U itself. Moreover, it still works if at each point on the boundary either U or $\nabla U \cdot \mathbf{n}$ is specified. The case where U is given on the boundary is called "*Dirichlet boundary conditions*", and the case where $\nabla U \cdot \mathbf{n}$ is given is called "*Neumann boundary conditions*". If we only have Neumann conditions, our W above is still a constant but not necessarily zero, so the solution U is unique up to addition of any arbitrary constant. We will only deal with Dirichlet boundary conditions from here on, but you may meet the Neumann conditions in later courses.]

The virtue of this uniqueness theorem is that it gives us a licence to make whatever assumptions or guesses we like, provided we can justify them afterwards by showing both Laplace's equation and the boundary conditions are satisfied: if they are, the solution we found must be the right one, even if our method involved some educated guesses.

Having proved uniqueness, we now demonstrate how to actually find solutions of Laplace's equation in some simple situations. In general $\Phi(\mathbf{r})$ can depend on all three coordinates, but we will confine ourselves to cases depending on two of the three coordinates: we will study the three most common coordinate systems as before:

- In Cartesian coordinates, we will take $\Phi(x,y)$, so Φ does not depend on z.
- In cylindrical polar coordinates, we will take $U(\rho,\phi)$ so U does not depend on z again, and we relabel Φ to U to avoid confusion with the angle ϕ.
- In spherical polar coordinates, we will take $U(r,\theta)$, so U does not depend on ϕ and we have rotational symmetry around the z axis.

The first two of these cases provide us with a nice geometrical interpretation. For $\Phi(x,y)$ or $U(\rho,\phi)$, we can forget about the $z-$ coordinate: then things reduce to a two dimensional problem, and we have boundary conditions given on the edge(s) of a region, (say a rectangle or circle) and we have to solve for Φ or U **inside** the given region. Now imagine Φ as a varying height h. Solving Laplace's equation in 2D subject to boundary

conditions is like taking a rubber sheet with its edges stuck to a rigid frame at the boundary with a "warp" in the third dimension: the frame fixes the height at the boundary, while the rubber tries to minimize its total area, which is equivalent to solving Laplace's equation.

For spherical polars $U(r,\theta)$ though, the 2-D interpretation no longer applies because the sphere still lives in 3-D.

Note: A "physical" example in three dimensions is as follows: suppose we take a uniform solid object (of arbitrary shape), and attach a large number of tiny thermostat-controlled heater/coolers to the surface, and set all the thermostats to some smoothly-varying function on the surface. The temperature inside, $T(\mathbf{r})$, will obey the heat equation

$$\frac{\partial T}{\partial t} = \kappa \nabla^2 T \quad ,$$

with κ a constant and boundary conditions set by our thermostats. If we wait a long enough time so the temperature distribution inside converges to a steady state, the LHS above will then be zero, so then the temperature inside the solid will solve Laplace's equation, with the given surface settings as the boundary condition. (If our boundary condition is T =constant independent of position, we just get the obvious boring solution T = constant inside; but if the boundary settings vary around the surface, it becomes an interesting problem.)

The choice of coordinates will be adapted to the geometry of the domain of interest and its boundaries, which usually makes calculations easier. For rectangular boundaries we use Cartesians, for circles or cylinders we use cylindrical polars, and for spherical boundaries we use spherical polars. For example, one may need to calculate the electrostatic potential outside a charged sphere. This would be very messy in Cartesian coordinates, and is much simpler if we use spherical polar coordinates instead. (This was one of the main reasons for studying Chapter 5)

7.3 2-D solutions of Laplace's equation in Cartesian coordinates

We first develop a general method for finding solutions $\Phi = \Phi(x,y)$ to Laplace's equation inside a rectangular domain, with given boundary conditions for Φ on all four edges of the rectangle. In Cartesian coordinates, as we saw above, Laplace's equation is

$$\nabla^2\Phi = \nabla\cdot(\nabla\Phi) = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = 0 \quad . \tag{7.3}$$

and in two dimensions we just drop the last term.

We will now try looking for a solution of the form

$$\Phi(x,y) = X(x)Y(y).$$

where $X(x)$ is some function of x only, and $Y(y)$ is some function of y only. Such a solution is called a **separable solution**. We cannot justify this in advance, but if it works then the uniqueness theorem tells us we are OK. It is possible to prove that any solution can be written as a sum (possibly an infinite sum) of separable solutions, but this is beyond the scope of this course.

Substituting the above Φ into (7.3) gives

$$\frac{\mathrm{d}^2X}{\mathrm{d}x^2}Y + X\frac{\mathrm{d}^2Y}{\mathrm{d}y^2} = 0.$$

Dividing this by XY gives

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2} = -\frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}y^2}.$$

Now, the left-hand side is a function of x **only**, and the right-hand side is a function of y **only**. This can only be satisfied if both sides are the same unknown constant.

Note: to prove the constant, $X(x)$ and $Y(y)$ must satisfy the above at any x,y inside our rectangle: so consider the above equation along a line (x_0,y) with fixed $x = x_0$ and varying y. The LHS is fixed, so the RHS must therefore be independent of y, i.e. constant. The same argument with y_0 fixed and x varying shows the LHS is constant, and it must be the same constant.

Now we call that constant $-\lambda$ with the minus sign for convenience, and both sides above equal $-\lambda$. Thus we have

$$\frac{\mathrm{d}^2X}{\mathrm{d}x^2} + \lambda X = 0 \quad \text{and} \quad \frac{\mathrm{d}^2Y}{\mathrm{d}y^2} - \lambda Y = 0\,.$$

If $\lambda \neq 0$, these equations are the differential equations for trigonometric and hyperbolic functions, which we met in chapter 1, so we know their general solutions as follows:

If λ is positive, define $k = \sqrt{\lambda}$ and the solution is

$$X = A\cos kx + B\sin kx, \quad Y = C\cosh ky + D\sinh ky,$$

where A,B,C,D are any constants. Multiplying these together,

$$\Phi = (A\cos kx + B\sin kx)\,(C\cosh ky + D\sinh ky)\ . \tag{7.4}$$

If λ is negative, define $k = \sqrt{-\lambda}$ and then the solution is

$$X = \hat{A}\cosh kx + \hat{B}\sinh kx, \quad Y = \hat{C}\cos ky + \hat{D}\sin ky.$$

where $\hat{A},\hat{B},\hat{C},\hat{D}$ are different constants. Then

$$\Phi = \left(\hat{A}\cosh kx + \hat{B}\sinh kx\right)\left(\hat{C}\cos ky + \hat{D}\sin ky\right)\ . \tag{7.5}$$

Note: in each of these solutions there is usually one more constant than we really need. For example if in (7.4) $AC \neq 0$ we can write

$$\Phi = AC\,(\cos kx + B/A\sin kx)\,(\cosh ky + D/C\sinh ky)$$

using just three constants AC, B/A and D/C: this means that in examples, one of the four constants can usually be set to 1. One way to do this is to write (7.4) as

$$\Phi = L\sin(kx+M)\sinh(ky+N)$$

for some constants L, M, and N. Usually this works fine, except it does not cover the case where $D = 0$.

Finally, we need to deal separately with the case $\lambda = 0$:
that easily gives us solutions $X = A_0x + B_0$ and $Y = C_0y + D_0$ so

$$\Phi = (A_0x + B_0)(C_0y + D_0)\ ,$$

with more constants A_0, B_0, C_0 and D_0. It is usually convenient to multiply this out and re-write it as

$$\Phi = \alpha + \beta x + \gamma y + \delta xy \tag{7.6}$$

with $\alpha,\beta,\gamma,\delta$ as alternative constants.

Remember, from linearity, **any sum** of any of the above functions with any k and any constants is also a solution of Laplace's equation. So, if we are given a boundary condition, solving Laplace's equation basically

reduces to choosing a **"pick-and-mix"** of any sum of the general solutions in order to satisfy all the given boundary conditions: if we manage to do that, then we have solved the problem (and our solution is unique). If we are lucky, a particular one of the separable solutions will do this, as we see in the next example.

Example 7.1. Find the solution of

$$\nabla^2\Phi \equiv \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} = 0 \qquad (*)$$

inside the rectangle D: $0 \le x \le a$, $0 \le y \le b$, given boundary conditions $\Phi = 0$ on the three sides $x = 0$, $y = 0$ and $x = a$; and $\Phi = \sin(p\pi x/a)$ on $y = b$, for some integer p.

We note here that Φ is zero along three of the sides, and non-zero along the "top" side with $y = b$. Also since $\sin 0 = 0$ and $\sin(p\pi a/a) = 0$, Φ is zero at the points $(0,b)$ and (a,b) so the given boundary condition is continuous at the corners.

Can we satisfy the boundary conditions in this case with one of the separable solutions above ? We consider them one by one. Clearly (7.6) will not work since it doesn't contain a sin. The form (7.4) is more promising, since if we take that equation and choose

$$A = 0 \qquad B = 1 \qquad k = \frac{p\pi}{a}$$

in there, the first bracket becomes $1\sin(p\pi x/a)$ which is the function we want on the boundary $y = b$. Now we just need to choose C, D to make the second bracket in 7.4 equal zero on the side $y = 0$, and 1 on the side $y = b$; this gives us two simultaneous equations for C, D :

$$C\cosh 0 + D\sinh 0 = 0$$

$$C\cosh(n\pi b/a) + D\sinh(n\pi b/a) = 1 \ ,$$

and the first of these implies $C = 0$, then the second gives $D = 1/\sinh(n\pi b/a)$.

Finally putting the above A, B, C, D back into 7.4 gives us

$$\Phi(x,y) = \sin\frac{p\pi x}{a}\sinh\frac{p\pi y}{a}/\sinh\frac{p\pi b}{a} \quad ;$$

this satisfies all the boundary conditions and Laplace's equation, so it is the unique solution.

In the above Example, we chose a "sin" in the boundary condition to make it easy: but for more general boundary conditions, using just one separable solution will not work.

However, since Laplace's equation is linear, we can add together separable solutions to get a more general solution. In many cases, including the Cartesian one, it is possible to prove that every solution can be written as a sum of separable solutions (this is called completeness of the separable solutions).

In the Cartesian case we would need to introduce different values of A for each k etc., which we typically would denote A_k. Since k can take any value, the "sum" of separable solutions can in general become an integral[1] over k; but for the rectangular boundaries in the example above we will only need to take integer values of p, call it n, so the general solution of Laplace's equation inside the rectangle becomes

$$\begin{aligned}\Phi(x,y) \quad = \quad & \alpha + \beta x + \gamma y + \delta xy \\ & + \sum_{n=1}^{\infty}(A_n\cos n\pi x/a + B_n\sin n\pi x/a)(C_n\cosh n\pi y/a + D_n\sinh n\pi y/a) \\ & + \sum_{n=1}^{\infty}(a_n\cosh n\pi x/b + b_n\sinh n\pi x/b)(c_n\cos n\pi y/b + d_n\sin n\pi y/b)\end{aligned} \qquad (7.7)$$

[1] This leads to the use of Fourier transforms, which is the next step, beyond this course, in Fourier methods

We note that the $\sin n\pi x/a$ terms vanish at $x=0$ and $x=a$ so they will fit Dirichlet boundary conditions which are zero on those boundaries. If multiplied by a $\sinh n\pi y$ they also vanish on $y=0$ so are non-zero only on $y=b$: to get similar forms which are zero at $y=b$ and non-zero at $y=0$ we need to take a combination of $\sinh n\pi y$ and $\cosh n\pi y$ which is zero at $y=b$: using the addition formula, this will turn out to look like $\sinh n\pi(b-y)/a$.

(The $\cos n\pi x/a$ terms are not zero on the boundary, but have vanishing derivative $\mathbf{n}.\nabla\Phi = \partial\Phi/\partial x$ at $x=0$ and $x=a$, so they will fit Neumann boundary conditions which are zero on those boundaries. Since we will stick to Dirichlet problems as examples in this course, we will find we are using only the sine terms not cos terms in our solutions).

For the other two sides at $x=0$ and $x=a$, we just repeat the above interchanging $x \leftrightarrow y$ and $a \leftrightarrow b$: so a solution which is non-zero only on side $x=a$ will look like $\sin n\pi y/b \sinh n\pi x/b$, and a solution which is non-zero only on the side $x=0$ will look like $\sin n\pi y/b \sinh n\pi(a-x)/b$.

From these remarks, we can see that in order to fit general boundary conditions, we can solve it if we break our function on the boundary into a (possibly infinite) sum of sin/cos functions i.e. a Fourier series.

Now we look at a boundary condition with a general function on one side:

Example 7.2. Consider the previous example but with $\Phi = g(x)$ on side $y=b$ for some given $g(x)$, and $\Phi = 0$ on the other three sides of the rectangle.

We try a linear combination of solutions of the form found above (keeping the conditions derived from the other parts of the boundary):

$$\Phi(x,y) = \sum_{n=1}^{\infty} D_n \sinh\frac{n\pi y}{a}\sin\frac{n\pi x}{a}.$$

Each term on the RHS is automatically a solution of Laplace's equation and is zero on the other three sides, so we just need to choose a set of constants D_n's to make this match the given $g(x)$ along the side $y=b$.

Putting in $y=b$ above gives us

$$\Phi(x,b) = \sum_{n=1}^{\infty} D_n \sinh\frac{n\pi b}{a}\sin\frac{n\pi x}{a} \qquad = g(x).$$

here the D_n and the sinh don't depend on x so we can rewrite this as

$$\Phi(x,b) = \sum_{1}^{\infty} E_n \sin\frac{n\pi x}{a} \qquad = g(x) \qquad (*)$$

with $E_n \equiv D_n \sinh(n\pi b/a)$.

Finding the coefficients E_n in equation $(*)$ is a standard problem in (arbitrary range) Fourier series from the previous Chapter. The answer is

$$E_n = \frac{2}{a}\int_0^a g(x)\sin\frac{n\pi x}{a}\,\mathrm{d}x.$$

Now we just need to evaluate this integral for all n, and then plug in $D_n = E_n/\sinh(n\pi b/a)$ back to the original equation to give us a solution

$$\Phi(x,y) = \sum_{n=1}^{\infty} \frac{E_n}{\sinh(n\pi b/a)}\sinh\frac{n\pi y}{a}\sin\frac{n\pi x}{a}.$$

By uniqueness, we have found *the* solution.

We still have a couple more issues to deal with. So far, we have seen how to solve the problem as a Fourier series when the boundary conditions are zero on three sides and non-zero on any one side.

If the boundary conditions are non-zero on all four sides but still zero at all four corners, we can solve this just by breaking it into four separate problems, each of which has non-zero boundary values on exactly one side: this gives four solutions $\Phi_1, \Phi_2, \Phi_3, \Phi_4$ each solving one different side: then add the four solutions, by linearity of Laplace's equation.

If the four corners are all one constant value, just subtract this constant from the boundary conditions, solve as above, and add the constant back to the final solution.

Finally, we have to deal with the case where the given boundary conditions are different (but still continuous) at the four corners. This can be dealt with by Eq. 7.6 above: given the four values at the corners, it is straightforward to choose our four constants $\alpha, \beta, \gamma, \delta$ to give a solution (call it Φ_0) which matches the given boundary values at all four corners, by starting with the $(0,0)$ corner, then the $(0,a)$, etc. Next, we subtract that $\Phi_0(x,y)$ from **all** the given boundary conditions on the edges to get a new set of boundary conditions for $\Phi_1 + \Phi_2 + \Phi_3 + \Phi_4$; solve Φ_1 to Φ_4 by treating the four sides separately as above: and finally add all five solutions $\Phi_0 + \ldots + \Phi_4$ to get the answer.

This whole process is quite lengthy, but we have seen how to do it in principle.

Example 7.3. Consider a rectangle with $0 \le x \le 2$, $0 \le y \le 1$, and boundary values for $\Phi(x,0) = \sin \pi x$ etc. as shown at the left diagram in Figure 7.1.

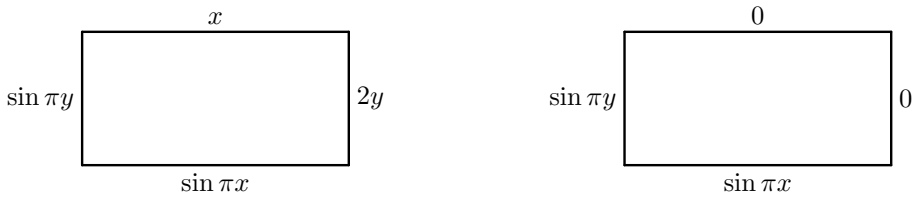

Figure 7.1: Left: boundary conditions on $\Phi(x,y)$. Right: boundary conditions after subtracting off $\Phi_0 = xy$ along the edges.

First we look at the boundary values at the four corners: reading these off the diagram we have $\Phi = 0$ at three corners and $\Phi(2,1) = 2$ at the corner $(x = 2, y = 1)$.
So, now we solve for the coefficients in $\Phi_0(x,y) = \alpha + \beta x + \gamma y + \delta xy$ so as to fit the given boundary values only at the four corners: starting at the origin and working out is easiest, so
$\Phi_0(0,0) = 0 \Rightarrow \alpha = 0$,
$\Phi_0(2,0) = 0 \Rightarrow \beta = 0$,
$\Phi_0(0,1) = 0 \Rightarrow \gamma = 0$,
$\Phi_0(2,1) = 2 \Rightarrow \delta = 1$;
therefore

$$\Phi_0(x,y) = 0 + 0 + 0 + 1xy \quad = xy$$

Now we evaluate Φ_0 along all four edges: it is zero on the left and bottom edges, it is $\Phi_0(2,y) = 2y$ on the right edge and $\Phi_0(x,1) = x$ along the top edge. Subtracting those from the original boundary conditions leaves the new boundary conditions in the right panel of Figure 7.1: by construction, these are zero at all corners. We can now match $\Phi(x,0)$ along the bottom side using

$$\Phi_1(x,y) = \frac{\sinh(\pi(1-y)) \sin \pi x}{\sinh(\pi)}$$

(this is like example 7.1), and match $\Phi(0,y)$ along the left-hand side with

$$\Phi_2(x,y) = \frac{\sinh(\pi(2-x))\sin\pi y}{\sinh(2\pi)} .$$

The full solution is

$$\Phi(x,y) = \Phi_0 + \Phi_1 + \Phi_2 .$$

Exercise 7.1. Find $\Phi(x,y)$ in $0 < x < \pi$, $0 < y < 1$, satisfying the following conditions:

$$\nabla^2\Phi = 0 \quad \text{in } 0 < x < \pi, \ \ 0 < y < 1,$$

$$\Phi = \sin x \text{ on } y = 0$$

and $\Phi = 0$ on the other three sides of the rectangle. Is the solution unique? □

7.4 2-D solutions of Laplace's equation in cylindrical polar coordinates

We now look at cylindrical polar coordinates: this is the natural choice where the boundary conditions are given on a circle in 2D or a cylinder in 3D. It will turn out a bit simpler than Cartesians, since there are no corners to worry about on the boundary.

We also change our label for our scalar field from Φ to U, to avoid confusion with the angle ϕ (of course, this is just a re-labelling and makes no real difference).

From chapter 5, in cylindrical polar coordinates (ρ,ϕ,z), the grad of a scalar field U is

$$\nabla U = \frac{\partial U}{\partial \rho}\mathbf{e}_\rho + \frac{1}{\rho}\frac{\partial U}{\partial \phi}\mathbf{e}_\phi + \frac{\partial U}{\partial z}\mathbf{e}_z$$

and the divergence of $\mathbf{F} = F_\rho\mathbf{e}_\rho + F_\phi\mathbf{e}_\phi + F_z\mathbf{e}_z$ is

$$\nabla\cdot\mathbf{F} = \frac{1}{\rho}\left[\frac{\partial(\rho F_\rho)}{\partial\rho} + \frac{\partial F_\phi}{\partial\phi} + \frac{\partial(\rho F_z)}{\partial z}\right].$$

Putting these together we obtain

$$\nabla^2 U \ \equiv\ \mathrm{div}(\nabla U) \ = \frac{1}{\rho}\left[\frac{\partial}{\partial\rho}\left(\rho\frac{\partial U}{\partial\rho}\right) + \frac{\partial}{\partial\phi}\left(\frac{1}{\rho}\frac{\partial U}{\partial\phi}\right) + \frac{\partial}{\partial z}\left(\rho\frac{\partial U}{\partial z}\right)\right] ,$$

which simplifies to

$$\nabla^2 U = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial U}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 U}{\partial\phi^2} + \frac{\partial^2 U}{\partial z^2} .$$

Consider the case when everything in the problem is independent of z, so $U = U(\rho,\phi)$. Once again we seek a **separable** solution, this time we will write it as

$$U(\rho,\phi) = R(\rho)S(\phi) .$$

where R and S are functions to be found. Putting this into $\nabla^2 U$, working out and dividing by RS gives

$$\frac{\rho}{R}\frac{\mathrm{d}}{\mathrm{d}\rho}\left(\rho\frac{\mathrm{d}R}{\mathrm{d}\rho}\right) = \frac{-1}{S}\frac{\mathrm{d}^2 S}{\mathrm{d}\phi^2} \qquad (*)$$

Once again, the LHS is a function of only ρ and the RHS is a function of only ϕ, so by the same argument as before, both sides are some (unknown) constant, call it λ.

Setting the RHS above to λ, the differential equation for $S(\phi)$ is then

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\phi^2} + \lambda S = 0 \,,$$

which we met before: if $\lambda > 0$, it has the general solution

$$S(\phi) = A\cos(\sqrt{\lambda}\phi) + B\sin(\sqrt{\lambda}\phi) \,.$$

If $\lambda < 0$ we would similarly have

$$S(\phi) = \hat{A}\cosh(\sqrt{-\lambda}\phi) + \hat{B}\sinh(\sqrt{-\lambda}\phi) \,.$$

But, in polar coordinates $S(\phi)$ must be periodic, i.e. the solution must be the same if we add 2π to ϕ, since any pair of values ϕ_0 and $\phi_0 + 2\pi$ represent the same point in space; the sinh and cosh solutions with $\lambda < 0$ cannot obey this, so are "forbidden" and we discard them. The sin and cos solutions will obey this periodic condition iff $\sqrt{\lambda}$ is an **integer**. Thus, the only allowed values of λ are $\lambda = m^2$ where m is a positive integer (without loss of generality) and we can write the solution for a particular integer m as

$$S(\phi) = A_m \cos m\phi + B_m \sin m\phi \,.$$

Now, going back to $R(\rho)$ and setting the LHS of (*) equal to $\lambda = m^2$ gives

$$\rho \frac{\mathrm{d}}{\mathrm{d}\rho}\left(\rho \frac{\mathrm{d}R}{\mathrm{d}\rho}\right) = m^2 R.$$

We guess a power-law solution $R = C\rho^q$ for constants C, q; substituting and working through, that simplifies to

$$q^2 = m^2$$

so $q = \pm m$. This is two independent solutions for q, and each has its own constant, so we write

$$R(\rho) = C_m \rho^m + D_m \rho^{-m} \,,$$

and again C_m, D_m are constants; finally multiplying out S and R, we have a solution for U of the form

$$U(\rho,\phi) = (A_m \cos m\phi + B_m \sin m\phi)\left(C_m \rho^m + D_m \rho^{-m}\right).$$

for any integer $m > 0$.

The case $\lambda = 0$ is again a special case: then we integrate twice giving $R = C_0 \ln\rho + D_0$, and $S = A_0\phi + B_0$. In most cases we set $A_0 = 0$ by requiring uniqueness on adding 2π to ϕ; (but note there are special cases where it is acceptable for U not to be unique, provided ∇U is unique. This happens in fluid dynamics, for example, where we are interested in the fluid velocity $\mathbf{v} = \nabla U$ rather than the potential U itself. In that case we require that ∇U be single valued, which allows us to use an A_0 term).

Combining the above, the general solution of Laplace's equation in cylindrical polars is a linear combination of all these above for the $m = 0$ case and every positive $m \geq 1$: each of these m has its own constants, so we get

$$U(\rho,\phi) = (A_0\phi + B_0)(C_0 \ln\rho + D_0) + \sum_{m=1}^{\infty} (A_m \cos m\phi + B_m \sin m\phi)\left(C_m \rho^m + D_m \rho^{-m}\right). \qquad (7.8)$$

Note that this form implies that boundary conditions like $g(R,\phi)$ given on a circle or cylinder of fixed $\rho = R$ leads to a Fourier series problem in ϕ (once the terms in $A_0\phi$ have been found). However, in many cases we

need only a finite number of terms and can use intelligent guesswork (essentially, including only terms with the **same** values of m which appear in the boundary conditions) to find the required set of constants.

Also, note the presence of both positive and negative powers of ρ on the RHS: if we are solving a problem **inside** a circle with boundary condition given on the circle, we will require all D_m to be zero for $m \geq 1$ so that the solution is bounded at the origin $\rho = 0$. Alternatively, we can be given boundary conditions on a circle and asked for a solution **outside** the circle, requiring the solution to be well-behaved at large $\rho \to \infty$: then we must set all C_m to zero for $m \geq 1$, and use only D_m terms.

Finally, we may be asked to solve Laplace's equation in an **annulus** between two circles of given radii, with boundary conditions given on both the inner and outer circles; in that case we will need to keep both C_m and D_m terms, and we'll get a pair of simultaneous equations for each m to match the given boundary conditions on both circles.

Example 7.4. Solve Laplace's equation $\nabla^2 U(\rho,\phi) = 0$ outside the unit circle, with boundary conditions $U(1,\phi) = 2\sin^2\phi$ on the unit circle, and $U \sim \ln\rho$ at large ρ.

First look at the general solution 7.8. That does not contain a $\sin^2$, but since $2\sin^2\phi = 1 - \cos 2\phi$, the latter form does look like a sum of two terms in 7.8: a constant ($m = 0$) term plus a $\cos 2\phi$ term which looks like an $m = 2$ term; so we can (correctly) guess that the same is true of the solution, i.e. we choose all $A_m \ldots D_m$ coefficients with $m = 1$ and $m \geq 3$ to be zero, so the infinite sum becomes just one term with $m = 2$. We also set $B_2 = 0$ since our boundary condition only contains a $\cos 2\phi$ not a $\sin 2\phi$.

The large-ρ condition implies $A_0 = 0$, and also $C_2 = 0$ since we don't want a ρ^{+2} term at large ρ.

Writing out 7.8 without all those zeros leaves us with our "educated guess" solution as

$$U = B_0 D_0 + B_0 C_0 \ln\rho + A_2 D_2 \cos(2\phi)\,\rho^{-2} .$$

Again this has several "redundant" constants, and we can just rewrite it as

$$U = \alpha + \beta \ln\rho + \gamma \cos(2\phi)\,\rho^{-2}$$

Finally, matching the given boundary values on the circle $\rho = 1$ gives us $\alpha = 1, \gamma = -1$, and the large ρ condition gives us $\beta = 1$, so the unique solution is

$$U(\rho,\phi) = 1 + \ln\rho - \frac{\cos 2\phi}{\rho^2} .$$

We can check this easily: it is a particular case of 7.8 so it does solve Laplace's equation. And it matches the given boundary conditions on $\rho = 1$ and large ρ; so it is the unique solution.

Exercise 7.2. Consider the region D defined by $a \leq \rho \leq b$, $0 \leq \phi \leq \pi$, $-\infty < z < \infty$. Sketch the region in a plane perpendicular to the z-axis which lies in D. On the boundaries $\rho = a$, $\phi = 0$ and $\phi = \pi$, $U = 0$ while on the boundary $\rho = b$, $U = \phi\sin\phi$. Find the solution U of Laplace's equation in D, independent of z, which satisfies these boundary conditions.

[You may assume that on $0 \leq \phi \leq \pi$

$$\phi\sin\phi = \sum_{k=1}^{\infty} \frac{16k}{\pi(4k^2-1)^2} \sin 2k\phi .]$$

□

Note: Finally, it is also worth noting that solutions like $\rho^n \cos(n\phi)$ can also be expanded as polynomials in x, y: for example, $\cos(4\phi) = 8\cos^4\phi - 8\cos^2\phi + 1$, and $\rho^4 = (x^2+y^2)^2$, therefore a bit of arithmetic leads to $\rho^4 \cos 4\phi \equiv x^4 - 6x^2y^2 + y^4$, and you can use the Cartesian formula to check that ∇^2 of that is zero. These may occasionally be useful, but they rapidly get unmanageable for large n.

7.5 Axisymmetric solutions of Laplace's equation in spherical polar coordinates

Now we consider what to do in problems with a naturally spherical geometry. First, we need to work out what $\nabla^2 U$ is in spherical polar coordinates.

As before, we have

$$\nabla^2 U \equiv \text{div}(\nabla U).$$

which is true in any coordinate system. Now in spherical polar coordinates,

$$\nabla U = \frac{\partial U}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial U}{\partial \theta}\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial U}{\partial \phi}\mathbf{e}_\phi$$

and the divergence of $\mathbf{F} = F_r\mathbf{e}_r + F_\theta\mathbf{e}_\theta + F_\phi\mathbf{e}_\phi$ is

$$\nabla\cdot\mathbf{F} = \frac{1}{r^2\sin\theta}\left[\frac{\partial(r^2\sin\theta F_r)}{\partial r} + \frac{\partial(r\sin\theta F_\theta)}{\partial\theta} + \frac{\partial(rF_\phi)}{\partial\phi}\right].$$

Putting these together we obtain

$$\nabla^2 U = \frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial U}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial U}{\partial\theta}\right) + \frac{\partial}{\partial\phi}\left(\frac{1}{\sin\theta}\frac{\partial U}{\partial\phi}\right)\right],$$

which simplifies to

$$\nabla^2 U = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial U}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial U}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 U}{\partial\phi^2}\right].$$

Many problems are *axisymmetric* – that is, there is no dependence on the ϕ coordinate. In such cases $U = U(r,\theta)$ and $\partial(\text{anything})/\partial\phi = 0$. As in the previous cases, we proceed by seeking a **separable** solution:

$$U(r,\theta) = R(r)S(\theta).$$

[different meanings from the R and S in the last section]. Thus $\nabla^2 U = 0$ becomes

$$\frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)S + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial S}{\partial\theta}\right)R\right] = 0$$

which rearranges to

$$\frac{1}{R(r)}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) = \frac{-1}{S(\theta)\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial S}{\partial\theta}\right).$$

Once again, the left-hand side is a function of r only, and the right-hand side is a function of θ only. But they are equal, and so they must both be some constant, say λ. Thus

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) - \lambda R = 0 \tag{7.9}$$

and

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}S}{\mathrm{d}\theta}\right) + \lambda S = 0. \tag{7.10}$$

We consider equation (7.10) first. If we define $w = \cos\theta$, then

$$\frac{\mathrm{d}}{\mathrm{d}w} = \frac{-1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta},$$

so equation (7.10) can be written in the form

$$\frac{\mathrm{d}}{\mathrm{d}w}\left((1-w^2)\frac{\mathrm{d}S}{\mathrm{d}w}\right)+\lambda S=0.$$

which is called **Legendre's differential equation**. We see in the next Section that only **Legendre polynomial** solutions $S=P_\ell(w)=P_\ell(\cos\theta)$ are allowed, i.e. the cases where $\lambda=\ell(\ell+1)$ and ℓ is an integer, and P_ℓ is the Legendre polynomial of order ℓ.

Going back to equation (7.9), inserting $\lambda=\ell(\ell+1)$ then $R(r)$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right)-\ell(\ell+1)R=0. \tag{7.11}$$

We try looking for a power-law solution, $R=Ar^p$ of this with A,p constant: inserting it we find

$$p(p+1)Ar^p=\ell(\ell+1)Ar^p$$

i.e. $p(p+1)=\ell(\ell+1)$. Given ℓ, this is a quadratic equation for p. It has solutions $p=\ell$ and $p=-(\ell+1)$. Hence the general solution for $R(r)$ is

$$R=Ar^\ell+\frac{B}{r^{\ell+1}}.$$

and so the solution for U is

$$U(r,\theta)=\left(Ar^\ell+\frac{B}{r^{\ell+1}}\right)P_\ell(\cos\theta).$$

Because ∇^2 is a linear operator, any linear combination of solutions is also a solution of Laplace's equation , so again the general solution is an infinite sum:

$$U(r,\theta)=\sum_{n=0}^{\infty}\left(A_nr^n+\frac{B_n}{r^{n+1}}\right)P_n(\cos\theta). \tag{7.12}$$

The individual functions on the right are **axisymmetric spherical harmonics** and they form a set of axisymmetric solutions of Laplace's equation which is complete, i.e. (7.12) can be shown to be the most general axisymmetric solution.

One can match arbitrary boundary conditions to an infinite series of Legendre polynomials using their orthogonality properties (see later). However, in this course we will stick to problems where only a few terms are needed and we can see what they are by intelligent guesswork: the essential rule is **only** to put into the prospective answer those Legendre polynomials which appear in the boundary conditions.

Example 7.5. A perfectly spherical conductor, centre 0, radius a, is placed in an otherwise uniform electric field $\mathbf{E}_0$. (Mathematically, the condition for a conductor is that the electrostatic potential U is constant.) What is the potential everywhere outside the conductor? And inside?

Outside the conductor ($r>a$), we want to solve $\nabla^2U=0$. The boundary conditions are that $U=$constant on $r=a$ and that far from the conductor $\nabla U\to\mathbf{E}_0$.

The unperturbed field (the one before the conductor was added) is $\mathbf{E}=E_0\mathbf{k}$, choosing the z-axis to align with the field. Converting this to the **e**'s of spherical polars, we have

$$\mathbf{E}_0=E_0\cos\theta\,\mathbf{e}_r-E_0\sin\theta\,\mathbf{e}_\theta$$

which is what the field must look like as $r\to\infty$: this has potential

$$U_0=E_0r\cos\theta+\text{ constant}=E_0rP_1(\cos\theta)+\text{ constant}.$$

(Note that this *is* a solution of Laplace's equation.) Now our potential

$$U = \sum_{n=0}^{\infty} \left(A_n r^n + \frac{B_n}{r^{n+1}} \right) P_n(\cos\theta) \;\to\; \sum_{n=0}^{\infty} A_n r^n P_n(\cos\theta)$$

as $r \to \infty$. But this must equal $U_0 = E_0 r P_1(\cos\theta)$+const. at large r, so we can deduce that $A_1 = E_0$, A_0 is an arbitrary constant, and $A_n = 0$ for all other n.

On $r = a$ we want U to be constant, i.e. it should not vary with θ. Now on $r = a$

$$U(a,\theta) = A_0 + \frac{B_0}{a} + \left(E_0 a + \frac{B_1}{a^2} \right) P_1(\cos\theta) + \sum_{n=2}^{\infty} \frac{B_n}{a^{n+1}} P_n(\cos\theta)$$

The potential on $r = a$ *will* vary with θ unless all the coefficients of $P_n(\cos\theta)$ ($n > 0$) each vanish. Hence we must have $B_1 = -E_0 a^3$ to make the bracket vanish, and $B_n = 0$ ($n \geq 2$). Hence finally the solution is

$$U(r,\theta) = A_0 + \frac{B_0}{r} + E_0 \left(r - \frac{a^3}{r^2} \right) \cos\theta.$$

Note that A_0 and B_0 are undetermined constants. To determine B_0 we need additional information to ascertain the potential difference between the surface of the conductor and a point at infinity. The constant A_0 will always be arbitrary, because the absolute value of the potential has no physical meaning (only its gradient is actually observable).

Inside, since U is constant on the boundary, it must be constant inside the conductor.

This last point has practical consequences. The voltage in space [in a static field] satisfies Laplace's equation. If you stand under an electricity pylon, there is a rather large voltage change—thousands of volts—between your head and your feet. But if you stand inside a wire cage (often called a Faraday cage), then the wire acts like a continuous conductor and equalizes the voltage over the cage and hence inside the cage too. That is why a wire cage provides a refuge from lightning. Cages also provide screening from electronic surveillance, or, by putting equipment inside them, safety for the people outside.

Exercise 7.3. Show that at a general point the following are solutions of Laplace's equation $\nabla^2 U = 0$.

1. $U = r^n \cos n\theta$, for an integer n, in cylindrical polar coordinates.
2. $U = r \sin\theta \cos\phi$, in spherical polar coordinates.

□

7.6 Introduction to Legendre polynomials

We now take a brief look at the Legendre polynomials. These are defined as the solutions of **Legendre's differential equation** which is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left((1-x^2) \frac{\mathrm{d}f}{\mathrm{d}x} \right) + \lambda f = 0.$$

or similar, where λ is an arbitrary constant. The solution of this is outside the scope of this course, but essentially we search for power-law solutions of the form

$$f(x) = \sum a_p x^p \quad .$$

Then, it can be shown that the series only converges at both $x = \pm 1$ if $\lambda = \ell(\ell+1)$ where ℓ is an integer, and we can take ℓ as a non-negative integer without loss of generality.

Then, the function $f(x)$ which satisfies the above D.E. for $\lambda = \ell(\ell+1)$ is called the **Legendre polynomial** of "degree" ℓ, usually written $P_\ell(x)$. (It is common to use letter ℓ for this integer, since when things are extended to 3-D **spherical harmonics**, letters n and m are generally used for other functions in the r and ϕ coordinates.)

There is an arbitrary multiplicative constant in each P_ℓ, which is chosen so that $P_\ell(1) = 1$ for all ℓ. It turns out that P_ℓ is an ℓ-th order polynomial, and involves only even/odd powers of w if ℓ is even/odd.

The solutions can be obtained by **Rodrigues' formula**

$$P_\ell(x) = \frac{1}{2^\ell \ell!} \frac{d^\ell}{dx^\ell}[(x^2-1)^\ell] \tag{7.13}$$

There is also a recurrence relation between them,

$$P_{\ell+1}(x) = \frac{1}{\ell+1}\left[(2\ell+1)xP_\ell(x) - \ell P_{\ell-1}(x)\right]$$

which gives all of them, working upwards from P_0 and P_1.

Starting from Rodrigues's formula

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \end{aligned}$$

then the recurrence relation gives subsequent ones as

$$\begin{aligned} P_2(x) &= \frac{1}{2}(3x^2-1) \\ P_3(x) &= \frac{1}{2}(5x^3-3x) \\ P_4(x) &= \frac{1}{8}(35x^4-30x^2+3) \\ &\text{etc} \end{aligned}$$

Another important property is **orthogonality**, i.e. the fact that

$$\begin{aligned} \int_{-1}^{1} P_m(w)P_n(w)\,dw &= 0 \quad \text{if } m \neq n \\ &= \frac{2}{2n+1} \quad \text{if } m = n \end{aligned}$$

This property enables us to express any general function as an infinite series of Legendre polynomials, by a device similar to that for calculating Fourier coefficients.

In this course we will only look at simple functions, in which case a general n-th order polynomial can be rearranged into a sum of the first n Legendre polynomials, e.g. suppose we are given a boundary condition in Laplace's equation looking like $f(w) = w^2 + w + 1$, ($w \equiv \cos\theta$) we need to choose a sum of Legendre polynomials to match this. We need $(2/3)P_2(w)$ to match the quadratic w^2 term. Then we need $1P_1(w)$ to match the linear term. Finally for the constant, the $(2/3)P_2$ has given us a $-1/3$ constant term, so to match the 1 we need $+(4/3)P_0$ on the right hand side. So in that example,

$$w^2 + w + 1 \equiv \frac{2}{3}P_2(w) + 1P_1(w) + \frac{4}{3}P_0(w)$$

and we can now put the right-hand-side into the general solution to Laplace's equation, 7.12 , and choose suitable constants to match.