1. Recap of prerequisites

Many formulas and ideas from earlier calculus courses are needed in this course. Here is a summary of the most important ones.

1.1 Trig functions: values. We can quickly estimate the value of a trigonometric function for any argument in $[0, 2\pi]$ by remembering a few things. First we have the table

$$\begin{array}{cccccc} & 0^\circ & 30^\circ & 45^\circ & 60^\circ & 90^\circ \\ \cos & 1 & \frac{\sqrt{3}}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} & 0 \\ \sin & 0 & \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{\sqrt{3}}{2} & 1 \end{array} \tag{1.1.1}$$

Then we remember that sin and cos swap when we shift by an odd multiple of 90°, they don't swap when we shift by a multiple of 180°. To get the sign we remember the table

$$\begin{array}{lccc} & \sin & \cos & \tan \\ 0^\circ - 90^\circ & + & + & + \\ 90^\circ - 180^\circ & + & - & - \\ 180^\circ - 270^\circ & - & - & + \\ 270^\circ - 360^\circ & - & + & - \end{array} \tag{1.1.2}$$

sometimes called the 'Add Sugar To Coffee' rule.

1.2 Example [Shifting to the first quadrant] Some applications of the above are

$$\begin{aligned} &\cos(-x) = \cos x, \quad \sin(-x) = -\sin x, \\ &\cos(\tfrac{\pi}{2} - x) = \sin x, \quad \sin(\tfrac{\pi}{2} - x) = \cos x, \\ &\cos(\pi - x) = -\cos x, \quad \sin(\pi - x) = \sin x. \end{aligned} \tag{1.2.1}$$

1.3 Trig functions: identities. The most important formulas to remember are

$$\begin{aligned} &\sin^2 A + \cos^2 A = 1 \\ &\cos(A + B) = \cos A \cos B - \sin A \sin B \\ &\sin(A + B) = \sin A \cos B + \cos A \sin B \end{aligned} \tag{1.3.1}$$

The double/half angle cases

$$\begin{aligned} &\sin 2x = 2 \sin x \cos x \\ &\cos 2x = \cos^2 x - \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x \\ &\cos^2 x = \tfrac{1}{2}(1 + \cos 2x) \\ &\sin^2 x = \tfrac{1}{2}(1 - \cos 2x) \end{aligned} \tag{1.3.2}$$

come up most often.

1.4 Example [More trig identities] Using the basic identities we can easily derive plenty more, such as

$$\begin{aligned} &\sec^2 A = 1 + \tan^2 A \\ &\cos A \, \cos B = \tfrac{1}{2}(\cos(A+B) + \cos(A-B)) \\ &\cos C + \cos D = 2 \, \cos \tfrac{1}{2}(C+D) \, \cos \tfrac{1}{2}(C-D) \end{aligned} \tag{1.4.1}$$

1.5 Example [Trig identities in integrals] A simple application of the AB formulas is

$$\int_0^{2\pi} \sin mx \cos nx \, dx = \tfrac{1}{2} \int_0^{2\pi} \big(\sin[(m+n)x] + \sin[(m-n)x] \big) \, dx = 0. \tag{1.5.1}$$

1.6 Taylor series. The most important Taylor series of all is

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \tag{1.6.1}$$

From this we can easily derive

$$\begin{aligned} \cosh x &\equiv \tfrac{1}{2} \left(e^x + e^{-x} \right) = 1 + \frac{x^2}{2!} + \ldots \\ \sinh x &\equiv \tfrac{1}{2} \left(e^x - e^{-x} \right) = x + \frac{x^3}{3!} + \ldots \end{aligned} \tag{1.6.2}$$

and also

$$\begin{aligned} \cos x &\equiv \tfrac{1}{2} \left(e^{ix} + e^{-ix} \right) = 1 - \frac{x^2}{2!} + \ldots \\ \sin x &\equiv \tfrac{1}{2i} \left(e^{ix} - e^{-ix} \right) = x - \frac{x^3}{3!} + \ldots \end{aligned} \tag{1.6.3}$$

We can get another useful Taylor series starting from the binomial expansion

$$\frac{1}{1+x} = x - x^2 + x^3 + \ldots \qquad |x| < 1. \tag{1.6.4}$$

Integrating term by term gives us

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots \tag{1.6.5}$$

1.7 Chain rule. This is about differentiating a function of a function. The simplest case is

$$\frac{d}{dx} F(y(x)) = \frac{dF}{dy} \frac{dy}{dx} \,. \tag{1.7.1}$$

More complicated is the case of several dependent variables x, y, z depending on one independent variable t:

$$\frac{d}{dt} F(x(t), y(t), z(t), t) = \frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} + \frac{\partial F}{\partial z}\frac{dz}{dt} + \frac{\partial F}{\partial t}\,. \tag{1.7.2}$$

Note the mixture of partial and total derivatives here.

1.8 Derivatives from parametric form. If we are given a function not as $y(x)$ but as

$$y = y(t), \quad x = x(t) \tag{1.8.1}$$

then the derivative is

$$\frac{dy}{dx} = \frac{(dy/dt)}{(dx/dt)} \tag{1.8.2}$$

1.9 Example [Cycloid] A point on the rim of a rolling wheel traces the curve

$$x = t - \sin t, \qquad y = 1 - \cos t \tag{1.9.1}$$

The derivative is (with the help of trigonometric identities)

$$\frac{dy}{dx} = \cot \tfrac{1}{2}t. \tag{1.9.2}$$

1.10 Example [A complicated differentiation] You may have come across the trick question "differentiate x^x wrt x".

Here we consider the more general example

$$y(x) = u(x)^{v(x)} \tag{1.10.1}$$

from which we have $\ln y = v \ln u$, hence $y'/y = (v/u)u' + v' \ln u$, hence

$$y' = u^v \left(\frac{v}{u}u' + v' \ln u\right). \tag{1.10.2}$$

Notice how

$$\begin{aligned} y = x^n &\Rightarrow y' = nx^{n-1} \\ y = a^x &\Rightarrow y' = a^x \ln a \end{aligned} \tag{1.10.3}$$

fall out as special cases of 1.10.2.

The chain rule lets us derive 1.10.2 more quickly.

1.11 Recursion relations in integrals. Let

$$I_n = \int_0^\infty x^n e^{-x}\, dx \tag{1.11.1}$$

where n is an integer ≥ 0. Integrating by parts leads to

$$I_n = nI_{n-1}. \tag{1.11.2}$$

This recursion relation has the obvious solution $I_n = n!$.

Recursion relations for integrals typically arise through integration by parts.

1.12 Rescaling. This refers to a simple variable changes of the type $x \to u = ax$ where a is a constant. It is very useful in simplifying integrals and differential equations. For example

$$\int_0^\infty F(ax)\,dx = \frac{1}{a}\int_0^\infty F(x)\,dx. \tag{1.12.1}$$

For a slightly more subtle example, consider the differential equation

$$\frac{d^2y}{dx^2} + y = 0 \tag{1.12.2}$$

which has the solution

$$y = A\sin(x+B). \tag{1.12.3}$$

By rescaling we can simplify

$$\frac{d^2y}{dt^2} + a^2 y = 0 \tag{1.12.4}$$

to the form (1.12.2) and hence derive the solution

$$y = A\sin(at+B). \tag{1.12.5}$$

***1.13 Improper integrals.** *Note that this is a starred section. Starred sections and examples are excluded from tests and Section A of the exam. But they are included for Section B of the exam and coursework.*

In the integral

$$\int_0^1 x^{\alpha-1}\,dx \tag{1.13.1}$$

if $0 < \alpha < 1$ the integrand is singular at $x = 0$ yet the integral can be evaluated. The integral

$$\int_1^\infty x^{-\alpha-1}\,dx \tag{1.13.2}$$

can be evaluated for $\alpha > 1$ even though the integration range is infinite.

These properties often help us test whether an integral converges even if we can't evaluate it.

For example, we cannot evaluate

$$\int_0^\infty \frac{\sin x\,dx}{1+x^2} \tag{1.13.3}$$

in terms of elementary functions but we know it converges, because

$$\int_0^\infty \frac{\sin x\,dx}{1+x^2} < \int_0^\infty \frac{|\sin x|\,dx}{1+x^2} < \int_0^\infty \frac{dx}{1+x^2} = \tfrac{\pi}{2} \tag{1.13.4}$$

Another example:

$$\int_0^\infty \sqrt{\frac{x^2+1}{x^5+1}}\,dx \tag{1.13.5}$$

converges, because for large x the integrand is $\propto x^{-3/2}$.

**1.14 Numerical integration. *Double-starred sections and examples are included for coursework only.*

Integration, we know, measures the area under a curve. We can make a rough estimate of that area by drawing a rectangle whose height is the mean height of the function. In other words,

$$\int_{x_0}^{x_0+\Delta x} y(x)\;dx \simeq \tfrac{1}{2}\Delta x(y_0+y_1) \tag{1.14.1}$$

where y_0 and y_1 are the integrand evaluated at the endpoints.

We can get improve accuracy by subdividing the integration range and taking N rectangles: In other words,

$$\frac{\Delta x}{2N}(y_0 + 2y_1 + \ldots + 2y_{N-1} + y_N). \tag{1.14.2}$$

We have $2y_1$ etc. here from two adjacent rectangles.

2. Fourier series

Fourier series are expansions in sines and cosines. They are especially important for studying oscillations and waves.

2.1 The main formula. Let $f(x)$ be a real-valued periodic function, with period 2π, meaning $f(x+2\pi) = f(x)$. Then $f(x)$ can be expressed as

$$\boxed{\begin{aligned} & f(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \\ & a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\, dx \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\, dx \end{aligned}} \tag{2.1.1}$$

and such series are known as Fourier series. We can verify (2.1.1) by working out the trigonometric integrals directly, using suitable trigonometric identities.

Some key theorems (which we will not prove) tell us that if $f(x)$ has a finite number of discontinuities and a finite number of extrema [the so-called Dirichlet conditions] then

(a) where $f(x)$ is continuous the Fourier series converges to the value of $f(x)$, and

(b) where $f(x)$ is discontinuous the Fourier series converges to the mean of the discontinuous values.

****2.2 Complex form.** For complex-valued $f(x)$ the formulas are actually more concise:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n\, e^{-inx} \qquad c_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)\, e^{inx}\, dx\ . \tag{2.2.1}$$

Here x is a real variable, while the c_n are complex constants.

To verify (2.2.1) we simply work out

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)\, e^{imx}\, dx = \sum_{n=-\infty}^{\infty} c_n \delta_{mn} = c_m\ . \tag{2.2.2}$$

As a special case, suppose $f(x)$ is real. Let us write

$$f(x) = \sum_{n=-\infty}^{\infty} c_n\, e^{-inx} = c_0 + \sum_{n=1}^{\infty}\left(c_n e^{-inx} + \bar{c}_n e^{inx}\right) + \sum_{n=1}^{\infty}(c_{-n} - \bar{c}_n)\, e^{inx}\ . \tag{2.2.3}$$

Here the LHS is real, and the first two expressions on the RHS are real too. So the last expression must be real too, and for all x. Hence $c_{-n} = \bar{c}_n$. Now let us separate the real and imaginary parts of c_n as

$$c_n = \tfrac{1}{2}(a_n + ib_n) \tag{2.2.4}$$

for $n >= 0$. But since $c_{-n} = \bar{c}_n$ we have

$$c_{-n} = \tfrac{1}{2}(a_n - ib_n). \qquad (2.2.5)$$

Substituting into (2.2.1) gives us a concise derivation of the main formula (2.1.1).

2.3 Example [Square wave] Consider

$$f(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 \text{ if } x > 0. \end{cases} \qquad (2.3.1)$$

in the domain $[-\pi, \pi]$ and periodic outside. This gives

$$a_0 = 1 \qquad a_{n>0} = 0 \qquad b_n = \frac{1 - \cos n\pi}{n\pi} \qquad (2.3.2)$$

or equivalently

$$f(x) = \tfrac{1}{2} + 2 \sum_{n \text{ odd}} \frac{\sin nx}{n\pi}\ . \qquad (2.3.3)$$

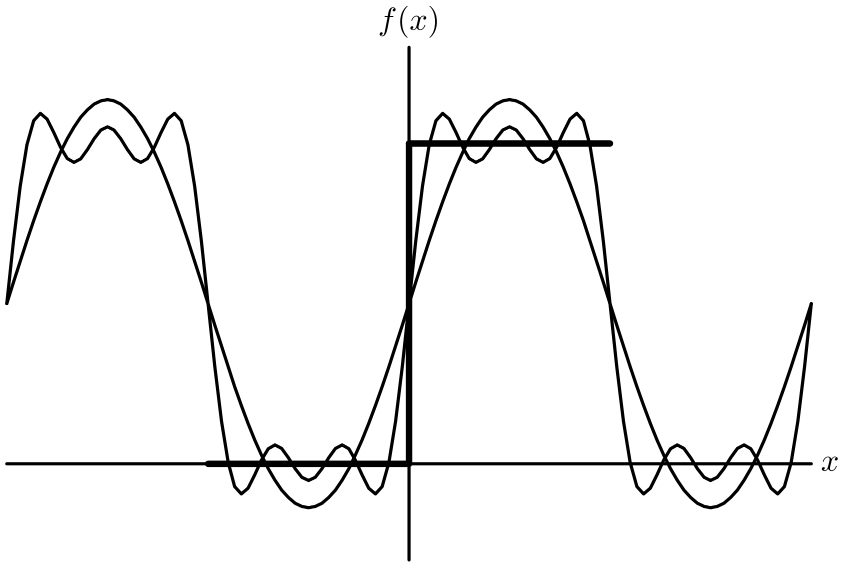

Figure 2.1: Square wave (as in equation 2.3.1 but with the vertical direction stretched for better visibility) and Fourier partial sums: two terms and four terms.

Figure 2.1 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) though we only considered the domain $[-\pi, \pi]$ the Fourier series automatically extends the domain to all real x by generating a periodic answer;

(iii) at discontinuities the Fourier series gives the mean value;

(iv) close to discontinuities the Fourier series overshoots.

The curious effect (iv) is called Gibbs' phenomenon—and adding more terms does not reduce the overshoot, it just moves the overshoot closer to the discontinuity.

Putting $x = -\frac{\pi}{2}$ in (2.3.3) gives an unexpected identity:

$$1 - \tfrac{1}{3} + \tfrac{1}{5} - \tfrac{1}{7} + \ldots = \frac{\pi}{4} \,. \tag{2.3.4}$$

2.4 Sine and cosine series. Often, the function we need to Fourier-expand will be either even or odd. In such cases, the coefficients simplify somewhat

$$\boxed{\begin{array}{llll} a_n = 0 & b_n = \dfrac{2}{\pi}\displaystyle\int_0^\pi f(x)\,\sin nx\,dx & & \text{if } f(x) \text{ is odd} \\ a_n = \dfrac{2}{\pi}\displaystyle\int_0^\pi f(x)\,\cos nx\,dx & & b_n = 0 & \text{if } f(x) \text{ is even} \end{array}} \tag{2.4.1}$$

and the series are known as Fourier sine and cosine series.

2.5 Example [Parabola] Consider $f(x) = x^2$ for x in $[-\pi, \pi]$ and periodic outside this domain. For $n > 0$ we write down a_n and integrate by parts, remembering that

$$\cos n\pi = (-)^n \tag{2.5.1}$$

we get

$$a_n = \frac{2}{\pi}\int_0^\pi x^2 \cos nx\,dx = \frac{4}{n^2}(-)^n. \tag{2.5.2}$$

We work out $a_0 = \frac{2}{3}\pi^2$ separately. Putting everything together we have

$$x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty} \frac{(-)^n}{n^2}\cos nx \,. \tag{2.5.3}$$

This may seem like a perverse way to calculate x^2. But put $x = 0$ and we get

$$\sum_{n=1}^{\infty} \frac{(-)^{n+1}}{n^2} = \frac{\pi^2}{12}. \tag{2.5.4}$$

And put $x = \pi$ and we get

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}. \tag{2.5.5}$$

2.6 Example [Manipulating sums] We can derive yet another series for π^2 from (2.5.5). To do this, we note that

$$\begin{aligned} \sum n^{-2} &= \sum_{n \text{ odd}} n^{-2} + \sum_{n \text{ even}} n^{-2} \\ \sum n^{-2} &= 4 \sum_{n \text{ even}} n^{-2} \end{aligned} \tag{2.6.1}$$

and solve to get

$$1 + \tfrac{1}{3^2} + \tfrac{1}{5^2} + \tfrac{1}{7^2} + \ldots = \frac{\pi^2}{8} \; . \tag{2.6.2}$$

2.7 Example [Rectifiers] Inside the power supply for an electrical appliance, the voltage may have the form

$$F(t) = |\sin t| \tag{2.7.1}$$

which is often called a full-wave rectifier. Here $V(t)$ will have a cosine series with

$$a_n = \frac{2}{\pi}\int_0^\pi \cos nt \, \sin t \, dt \; . \tag{2.7.2}$$

To do the integral we use the identity $2\cos nt \, \sin t = \sin(n+1)t - \sin(n-1)t$. The Fourier series is

$$F(t) = \frac{2}{\pi}\left[1 - \frac{2\cos nt}{n^2-1}\right] \; . \tag{2.7.3}$$

Also important is the half-wave rectifier, where the voltage has the form

$$H(t) = \begin{cases} \sin t & \text{if } 0 \le t \le \pi \\ 0 & \text{if } -\pi \le t \le 0 \; . \end{cases} \tag{2.7.4}$$

We can derive its Fourier series easily from $F(t)$ by noting that

$$H(t) = \tfrac{1}{2}(\sin t + |\sin t|) \; . \tag{2.7.5}$$

***2.8 Parseval's relation.** Fourier components satisfy the identity

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)\, dx = \tfrac{1}{2}a_0^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2) \tag{2.8.1}$$

which is known as Parseval's relation.

****2.9 Parseval's relation: concise derivation.** One can verify Parseval's relation by substituting the main formula and working through, but a more concise way is to start from the complex expansion (2.2.1), which leads immediately to

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\, dx = \sum_{n=-\infty}^{\infty} |c_n|^2 \; . \tag{2.9.1}$$

If $f(x)$ is real we write $c_n = \frac{1}{2}(a_n + ib_n)$ and $c_{-n} = \frac{1}{2}(a_n - ib_n)$ as before. Substituting and then simplifying we get (2.8.1).

***2.10 Example** [Parabola again] From section 2.5 we have

$$f(x) = x^2 \qquad a_0 = \frac{2\pi^2}{3} \qquad a_{n>0} = \frac{4}{n^2}(-1)^n \tag{2.10.1}$$

Substituting in Parseval's relation and simplifying gives

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90} \,. \tag{2.10.2}$$

****2.11 Orthogonal functions.** The functions $\sin nx$ and $\cos nx$ are examples of what are called orthogonal functions. In general, orthogonal functions are any set of functions $F_n(x)$ satisfying

$$\int F_m(x)\, F_n(x)\, dx = N_{mn}\delta_{mn} \tag{2.11.1}$$

with N_{mn} constant. If $N_{mn} = 1$ we call the functions orthonormal. (We can always arrange that by redefining $F_n(x)$.)

If the $F_n(x)$ are orthonormal, then we may have series expansions of the form

$$f(x) = \sum_n c_n\, F_n(x) \,. \tag{2.11.2}$$

It is not automatic that such a series will converge, but for well-behaved functions it usually does. Assuming the series does converge, it follows from orthonormality that the coefficients are given by

$$c_n = \int f(x)\, F_n(x)\, dx. \tag{2.11.3}$$

There is another interpretation of such series. Let us look for a series of the type $\sum_n c_n F_n(x)$ such that it best approximates $f(x)$, in the sense that the approximation error

$$E = \int \left(f(x) - \textstyle\sum_n c_n F_n(x)\right)^2 dx \tag{2.11.4}$$

is minimized. To minimize E we solve

$$\frac{\partial E}{\partial c_m} = 0 \tag{2.11.5}$$

which gives us

$$\int F_m(x)\left(f(x) - \textstyle\sum_n c_n F_n(x)\right) dx = 0 \tag{2.11.6}$$

or

$$\int f(x)\, F_m(x)\, dx = \textstyle\sum_n \displaystyle\int F_m(x) F_n(x)\, dx = c_m \tag{2.11.7}$$

which is the same answer as (2.11.3).

****2.12 Another variable.** We have so far assumed the Fourier coefficients are constants. But we are free to make them functions of another independent variable. That is, we can have

$$f(x,t) = \tfrac{1}{2} a_0(t) + \sum_{n=1}^{\infty} \left(a_n(t) \cos nx + b_n(t) \sin nx\right). \tag{2.12.1}$$

****2.13 Example** [Diffusion equation] Now suppose $f(x,t)$ satisfies the partial differential equation

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial f}{\partial t}\,. \tag{2.13.1}$$

This equation implies the ordinary differential equations

$$-n^2 a_n(t) = \frac{da_n}{dt} \qquad -n^2 b_n(t) = \frac{db_n}{dt} \tag{2.13.2}$$

which are easily solved to give

$$a_n(t) = e^{-n^2 t} a_n(0) \qquad b_n(t) = e^{-n^2 t} b_n(0). \tag{2.13.3}$$

In this solution the structure present in $f(x,0)$ gradually fades away, leaving just $f(x,t) = \frac{1}{2}a_0$ for large t. Equation (2.13.1) is the simplest form of the so-called diffusion equation. When Fourier invented his series, he was interested in studying the diffusion of heat.

3. Calculus of Variations

The calculus of variations is about finding paths such that some integral along the path is extremized.

3.1 Statement of the problem. We want to find a path $y(x)$ connecting two given points (x_1, y_1) and (x_2, y_2) such that a given integral

$$\int_{x_1}^{x_2} L(y, y', x)\, dx \tag{3.1.1}$$

(here y' denotes dy/dx as usual) is extremized.[1] By this we mean that the integral along $y(x)$ equals the integral along an infinitesimally close path $y(x) + \delta y(x)$. The condition is written as

$$\delta \int_{x_1}^{x_2} L(y, y', x)\, dx = 0\ . \tag{3.1.2}$$

Note that $\delta y(x) = 0$ at the endpoints, since the latter are given.

Integral conditions of the type (3.1.2) are known as variational principles. The function $L(y, y', x)$ is usually called a Lagrangian.

3.2 The Euler-Lagrange equation. Remarkably, the integral condition (3.1.2) can be reduced to a differential equation

$$\boxed{\frac{\partial L}{\partial y} = \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)} \tag{3.2.1}$$

known as the Euler-Lagrange equation. Note how this equation involves both partial and total derivatives.

****3.3 Derivation of the Euler-Lagrange equation.** We can write the variational equation (3.1.2) as

$$\int_{x_1}^{x_2} \left(\frac{\partial L}{\partial y}\,\delta y + \frac{\partial L}{\partial y'}\,\delta y'\right) dx = 0 \tag{3.3.1}$$

Integrating the second term by parts and putting $\delta y = 0$ at the ends gives

$$\int_{x_1}^{x_2} \left(\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\right) \delta y\; dx = 0 \tag{3.3.2}$$

Since this must be true under arbitrary δy, the integrand must be zero, giving the Euler-Lagrange equation.

---

[1] We will use extremum to denote maximum, minimum, or inflection point. An alternative term is stationary.

3.4 Example [Dynamics revisited] In dynamics we are often interested in the solution $x(t)$ of a differential equation of the type

$$\frac{d^2x}{dt^2} = -\frac{\partial}{\partial x} V(x,t) \ . \tag{3.4.1}$$

In fact this differential equation is the Euler-Lagrange equation that results from extremizing

$$\int \left(\tfrac{1}{2} \left(\frac{dx}{dt} \right)^2 - V(y,t) \right) \, dt \tag{3.4.2}$$

as we can easily verify.

3.5 Example [Importance of the derivative] The Lagrangian $L(y,y',x)$ can be independent of y and x it it wants, but it *must* depend on y'. Otherwise, that is, if we attempt to extremize $\int L(y,x)\,dx$ we get $\partial L/\partial y = 0$ for the Euler-Lagrange equation, which is meaningless.

3.6 Example [Non-uniqueness of the Lagrangian] The curve $y(x)$ that extremizes

$$\int L(y,y',x)\,dx \tag{3.6.1}$$

between two fixed points will also extremize

$$\int \left(L(y,y',x) + \frac{dF}{dy}\,y' \right) \, dx \tag{3.6.2}$$

for any well-behaved $F(y)$.

To see this, we work out the terms introduced to the Euler-Lagrange equations $F(y)$. These are

$$\frac{d}{dx}\left(\frac{dF}{dy}\right) - \frac{d^2F}{dy^2}\,y' \tag{3.6.3}$$

and they cancel by the chain rule.

A more elegant argument is that the extra term $(dF/dy)\,y'$ is just dF/dx. This an exact differential, whose contribution to the integral is independent of path.

3.7 Arc length. Many problems in the calculus of variations involve arc length, meaning distance along a curve. Recall that the arc length of $y(x)$ is

$$\int \sqrt{1+y'^2}\,dx \tag{3.7.1}$$

whereas for a parametric curve $x(t)$, $y(t)$ the arc length is

$$\int \sqrt{\dot{x}^2+\dot{y}^2}\,dt \tag{3.7.2}$$

where $\dot{x}$ means dx/dt and $\dot{y}$ means dy/dt.

3.8 Example [Straight line] To get the shortest distance between two points (x_1, y_1) and (x_2, y_2) we extremize

$$\int_{x_1}^{x_2} \sqrt{1 + y'^2}\, dx \tag{3.8.1}$$

for which the Euler-Lagrange equations are

$$y'' = 0. \tag{3.8.2}$$

The solution is obviously a straight line passing through the given points.

3.9 The Hamiltonian. The Euler-Lagrangian equations are equivalent to

$$\frac{\partial L}{\partial x} - \frac{d}{dx}\left(L - y' \frac{\partial L}{\partial y'} \right) = 0 \tag{3.9.1}$$

as we can verify using the chain rule. This means that if the Lagrangian has no *explicit* dependence on x then

$$\boxed{y' \frac{\partial L}{\partial y'} - L = \text{const.}} \tag{3.9.2}$$

The expression in (3.9.2) is called the Hamiltonian (and sometimes the first integral).

We see that if the Lagrangian does not depend on the independent variable then the Euler-Lagrange equation simplifies to a first-order equation.

3.10 Example [Fermat's principle] An important type of problem is when

$$\int_{x_1}^{x_2} \mu(y) \sqrt{1 + y'^2}\, dx \tag{3.10.1}$$

is to be extremized, $\mu(y)$ being a given well-behaved function. Working out the Hamiltonian, we see

$$\frac{\mu(y)}{\sqrt{1 + y'^2}} = \text{const.} \tag{3.10.2}$$

This type of problem has a special interpretation for light rays. Fermat's principle states that light rays take paths that make the travel time extremal (usually a minimum, but not always). If $\mu(y)$ is the refractive index (i.e., inverse speed of light) then (3.10.1) represents the light travel time, and making the integral extremal amounts to Fermat's principle.

3.11 Example [Fermat's principle and Snell's law] Consider a light ray travelling through two substances of different refractive index μ_1 and μ_2 as depicted in Figure ??.

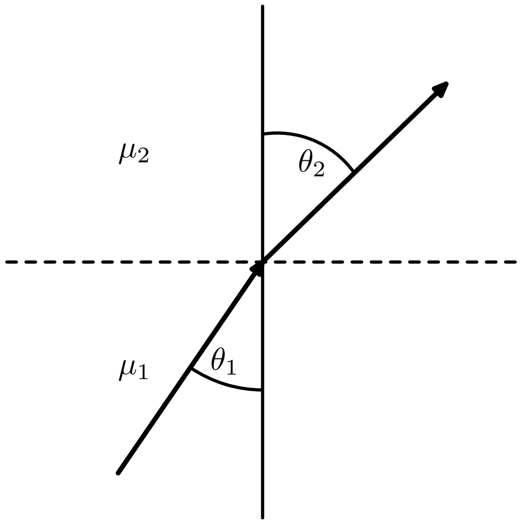

Figure 3.1: Illustration of Snell's law.

Taking x as the horizontal coordinate and y as the vertical coordinate, we immediately apply Fermat's principle and hence get (3.10.2).

In the figure, $y' = \cot\theta$ for the straight paths. Using $\csc^2\theta = 1 + \cot^2\theta$ gives us

$$\frac{1}{\sqrt{1+y'^2}} = \sin\theta\ . \tag{3.11.1}$$

and hence

$$\mu(y)\sin\theta = \text{const.} \tag{3.11.2}$$

The constant in the μ_1 and μ_2 parts must be the same, and hence

$$\mu_1 \sin\theta_1 = \mu_2 \sin\theta_2. \tag{3.11.3}$$

This relation is known in optics as Snell's law.

3.12 Example [Arches and catenaries] If we have a flexible cable and fix its ends at (x_1, y_1) and (x_2, y_2) then

$$\int_{x_1}^{x_2} y\sqrt{1+y'^2}\,dx \tag{3.12.1}$$

has the interpretation of potential energy. The integral having the Fermat-principle form (3.10.1), we have

$$\frac{y}{\sqrt{1+y'^2}} = \text{const.} \tag{3.12.2}$$

The solution is

$$y = c_1\cosh(x/c_1 + c_2)\,. \tag{3.12.3}$$

The curve is called a catenary. If y is measured upwards, we have a minimum of potential energy and the shape of the suspended cable. If y is measured downwards we have a maximum of potential energy and the shape of an arch.

***3.13 Example** [Great circles] On a sphere (of unit radius, say), instead of x, y we consider the angles θ, ϕ. Here ϕ is the longitude and θ is $90° -$ latitude.

The arc length along a curve $\theta(\phi)$ is[2]

$$\int \sqrt{\theta'^2 + \sin^2\theta}\, d\phi \tag{3.13.1}$$

where θ' means $d\theta/d\phi$. Extremizing the integral leads to

$$\frac{\sin^2\theta}{\sqrt{\theta'^2 + \sin^2\theta}} = \text{const } 1/\kappa \tag{3.13.2}$$

or equivalently

$$\theta'^2 + \sin^2\theta = \kappa^2 \sin^4\theta\ . \tag{3.13.3}$$

This differential equation cannot be solved in terms of elementary functions, but we can infer something about the solutions. Consider a curve segment where θ increases. Along this segment, $\sin^4\theta$ will increase by a larger factor than $\sin^2\theta$ and meanwhile, to satisfy (3.13.3), θ'^2 will have to increase. As a result the extremal curves $\theta(\phi)$ get steeper for larger ϕ.

In fact the solutions of (3.13.3) are sections of the sphere with the same diameter as the sphere, and are known as great circles. Airline routes are great circles, and have strange shapes on ordinary maps. In particular, a plane from London to Tokyo (which is south of London) will first head north.

****3.14 Variation with constraints.** We already have Lagrangians and Euler-Lagrange equations; we are about to see yet another of Lagrange's ideas, the multipliers.

Let us extremize

$$\int_{x_1}^{x_2} \Big[L(y, y', x) + \lambda F(y, y', x) \Big]\, dx \tag{3.14.1}$$

where λ is a parameter. We will obtain a family of curves $y(x, \lambda)$. Let us then choose λ so that

$$\int_{x_1}^{x_2} F(y, y', x)\, dx = 1 \text{ (say)}. \tag{3.14.2}$$

This procedure extremizes the original integal (3.1.1) subject to the constraint (3.14.2).

****3.15 Example** [The cable, again] Let us consider the suspended cable again, but now fix its length to be 1. We now have to solve

$$\delta \int_{x_1}^{x_2} y\sqrt{1 + y'^2}\, dx = 0 \tag{3.15.1}$$

[2] We will derive this in a later chapter.

subject to

$$\int_{x_1}^{x_2} \sqrt{1+y'^2}\, dx = 1\ . \tag{3.15.2}$$

This problem is equivalent to extremizing

$$\int_{x_1}^{x_2} (y+\lambda)\sqrt{1+y'^2}\, dx\ . \tag{3.15.3}$$

and then setting λ so as to satisfy the length constraint. The Hamiltonian is

$$\frac{y+\lambda}{\sqrt{1+y'^2}} = \text{const} \tag{3.15.4}$$

which is the same as (3.12.2) except for a constant shift in y.

In this case, the constraint makes only a trivial difference to the equations.

3.16 Example [Areas and perimeters] Suppose we have a long straight wall and want to put down some fencing of fixed total length so as to enclose the maximum area between the wall and the fence.

Taking the wall to be the x axis and $y(x)$ to be the shape of the fence, we have to extremize

$$\int_{x_1}^{x_2} (y+\lambda\sqrt{1+y'^2})\, dx \tag{3.16.1}$$

and then set λ to give the correct length of fence. Working out the Hamiltonian gives us

$$y\sqrt{1+y'^2} = a \tag{3.16.2}$$

where λ has been absorbed inside the constant a.

Equation (3.16.2) is solved by

$$x = a\cos\phi + b, \qquad y = a\sin\phi \tag{3.16.3}$$

in other words, a semicircle.

A corollary of this result is that, of all closed curves with a given perimeter, a circle encloses the largest area.

4. Vector differentiation

In this chapter we will introduce the idea of vector fields, and the different differential operators (gradient, divergence, curl) that they bring up.

4.1 Vectors and vector fields. *Vectors* are an abstraction of physical concepts like displacement and force, which have magnitude and direction, and are additive. A *vector field* is a function of spatial position whose values are vectors.

4.2 Cartesian coordinates. In 3D, we commonly use right-handed cartesian coordinates (x, y, z). We write $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ for unit vectors the along x, y, z directions. See Figure ??.

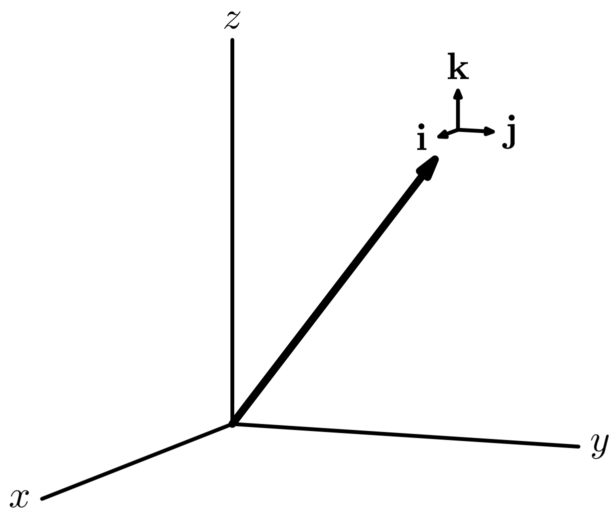

Figure 4.1: Cartesian coordinates and unit vectors.

4.3 Position vectors. Spatial displacement in three dimensions is a vector. In particular, the displacement from the origin of coordinates to a point (x, y, z) is a vector. This is written as

$$\boxed{\mathbf{r} \equiv x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}} \tag{4.3.1}$$

and called the position vector of the point (x, y, z). We will often refer simply to the 'point $\mathbf{r}$', since there is no ambiguity in doing so.

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

4.4 Lines. An expression of the type

$$\mathbf{r} = \mathbf{p} + s\mathbf{q}, \qquad -\infty < s < \infty \tag{4.4.1}$$

denotes a line through $\mathbf{p}$ in direction $\mathbf{q}$. In particular $\mathbf{r} = s\,\mathbf{k}$, $-\infty < s < \infty$ is the z axis, while a line going through $\mathbf{r}_1$ and $\mathbf{r}_2$ is $\mathbf{r} = \mathbf{r}_1 + s(\mathbf{r}_2 - \mathbf{r}_1)$.

4.5 Example [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 last point and $\mathbf{a}$ is

$$\mathbf{r} = \mathbf{a} + s(\tfrac{1}{2}\mathbf{b} + \tfrac{1}{2}\mathbf{c} - \mathbf{a}), \qquad -\infty < s < \infty \tag{4.5.1}$$

and is called the median through $\mathbf{a}$. Putting $s = \frac{2}{3}$ 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 medians of a triangle intersect.

4.6 Vector notation. Vectors are commonly written like $\mathbf{v}$ in print and $\underline{v}$ or $\vec{v}$ in handwriting. Magnitude of vectors are written like $|\mathbf{v}|$ (or $|\underline{v}|$ or $|\vec{v}|$) or simply v.

Sometime we use explicit component notation:

$$(v_x, v_y, v_z) \equiv v_x\,\mathbf{i} + v_y\,\mathbf{j} + v_z\,\mathbf{k} \tag{4.6.1}$$

For a unit vector along $\mathbf{v}$ we use $\hat{\mathbf{v}}$ (or $\hat{v}$ in handwriting). Thus

$$\hat{\mathbf{v}} \equiv \mathbf{v}/v \tag{4.6.2}$$

4.7 Dot product. This is defined through

$$\begin{aligned} &\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1 \\ &\mathbf{i}\cdot\mathbf{j} = 0, \quad \text{etc} \end{aligned} \tag{4.7.1}$$

and we interpret

$$\mathbf{u}\cdot\mathbf{v} \quad \text{as} \quad uv\cos\langle\text{mutual angle}\rangle\,. \tag{4.7.2}$$

4.8 Cross product. This is defined through

$$\mathbf{i}\times\mathbf{j} = -\mathbf{j}\times\mathbf{i} = \mathbf{k}, \quad \mathbf{j}\times\mathbf{k} = -\mathbf{k}\times\mathbf{j} = \mathbf{i}, \quad \mathbf{k}\times\mathbf{i} = -\mathbf{i}\times\mathbf{k} = \mathbf{k} \tag{4.8.1}$$

and we interpret

$$\mathbf{u}\times\mathbf{v} \quad \text{as} \quad uv\sin\langle\text{mutual angle}\rangle\,\langle\text{unit vector} \perp \mathbf{u},\mathbf{v}\rangle\,. \tag{4.8.2}$$

4.9 Vector identities. There are numerous vector identities. Two basic identities are

$$\boxed{\mathbf{v}\cdot\mathbf{v} = v^2 \qquad \mathbf{v}\times\mathbf{v} = 0} \tag{4.9.1}$$

while

$$\boxed{\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B})} \tag{4.9.2}$$

is another important one, which we will derive later.

4.10 Example [Diagonals of a cube] The position vectors

$$l\,\mathbf{i} + m\,\mathbf{j} + n\,\mathbf{k}, \qquad l, m, n = 0, 1 \tag{4.10.1}$$

are the vertices of a cube, while

$$(\,\mathbf{i} \pm \mathbf{j} \pm \mathbf{k})/\sqrt{3} \tag{4.10.2}$$

are units vectors along its diagonals.

Taking dot products, the angle between diagonals is $\arccos\left(\pm\frac{1}{3}\right)$ amounting to $\simeq 71°/109°$.

4.11 Surfaces. We will often come across expressions of the type

$$\Phi(\mathbf{r}) = 0. \tag{4.11.1}$$

These are surfaces. A simple example is $r^2 = 1$, which is a sphere of unit radius centred at the origin.

4.12 Example [A sphere] The geometrical interpretation of

$$|\mathbf{r} - \mathbf{k}| = 1 \tag{4.12.1}$$

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 $x^2 + y^2 + z^2 - 2z = 0$.

4.13 Functional dependence. A vector function of $\mathbf{r}$, as in

$$\mathbf{v}(\mathbf{r}) \tag{4.13.1}$$

is a vector field. A vector field may depend on additional variables (typically time) as well

$$\mathbf{v}(\mathbf{r}, t) \tag{4.13.2}$$

and sometimes we are interested in a vector field along a path

$$\mathbf{v}(\mathbf{r}(s), t). \tag{4.13.3}$$

Differentiation with respect to t is straightforward and derivatives of products are

$$\begin{aligned}\frac{\partial}{\partial t}(\mathbf{u}\cdot\mathbf{v}) &= \frac{\partial\mathbf{u}}{\partial t}\cdot\mathbf{v} + \mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial t}\\ \frac{\partial}{\partial t}(\mathbf{u}\times\mathbf{v}) &= \frac{\partial\mathbf{u}}{\partial t}\times\mathbf{v} + \mathbf{u}\times\frac{\partial\mathbf{v}}{\partial t}\,.\end{aligned} \tag{4.13.4}$$

as you'd expect.

But differentiation with respect to x, y, z does *not* behave in the same straightforward way. For those we need a new ingredient, the ∇ operator.

4.14 The del operator. Del (also called nabla)

$$\boxed{\nabla \equiv \left(\mathbf{i}\,\frac{\partial}{\partial x} + \mathbf{j}\,\frac{\partial}{\partial y} + \mathbf{k}\,\frac{\partial}{\partial z}\right)} \tag{4.14.1}$$

is a vector differential operator. It is not a vector but shares some properties with vectors, hence the notation.

4.15 Gradient. The gradient of a scalar field $\nabla\Phi(\mathbf{r})$ is a vector field. The vector represents the direction and magnitude of the maximum rate of change.

For a geometrical interpretation refer to Figure ??. The equation $\Phi(\mathbf{r}) = \text{const}$ represents a surface (curve in 2D) and $\nabla\Phi(\mathbf{r})$ is the local normal vector, while

$$\boxed{(\mathbf{r} - \mathbf{p}) \cdot \mathbf{n} = 0} \qquad \mathbf{n} \equiv \nabla\Phi\Big|_{\mathbf{r}=\mathbf{p}} \tag{4.15.1}$$

is the tangent plane passing through $\mathbf{p}$.

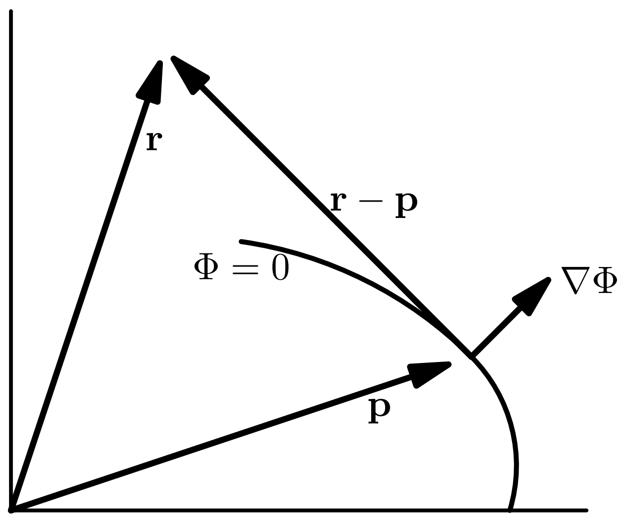

Figure 4.2: Geometrical interpretation of the gradient.

4.16 Example [A tangent plane] Consider the point $\mathbf{p} = (1, 1, 1)$ on the sphere $r^2 = 3$.

The unit normal to the surface at $\mathbf{p}$ is

$$\frac{\mathbf{i} + \mathbf{j} + \mathbf{k}}{\sqrt{3}} \tag{4.16.1}$$

and the tangent plane is

$$\big(\mathbf{r} - (\,\mathbf{i} + \mathbf{j} + \mathbf{k})\big) \cdot \frac{(\,\mathbf{i} + \mathbf{j} + \mathbf{k})}{\sqrt{3}} = 0 \tag{4.16.2}$$

which simplifies to

$$x + y + z = 3. \tag{4.16.3}$$

4.17 Example [Radial gradient] It is easily verified that $\nabla r = \hat{\mathbf{r}}$, from which

$$\nabla f(r) = \frac{df}{dr}\,\hat{\mathbf{r}} \tag{4.17.1}$$

and in particular

$$\nabla(r^n) = nr^{n-2}\mathbf{r}. \tag{4.17.2}$$

4.18 Example [Angular momentum] Some important problems in dynamics have the form

$$\ddot{\mathbf{r}} = -\nabla\Phi(r). \tag{4.18.1}$$

The rhs is called a central force, and implies (as we can easily verify)

$$\mathbf{r}\times\dot{\mathbf{r}} = \text{const} \tag{4.18.2}$$

which is called angular-momentum conservation.

4.19 Chain rule. We express this concisely as a vector operator

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \frac{d\mathbf{r}}{dt} \cdot \nabla \tag{4.19.1}$$

which can act on scalar fields $\Phi(\mathbf{r}(t), t)$ or vector fields $\mathbf{F}(\mathbf{r}(t), t)$.

4.20 Divergence and curl. The divergence $\nabla \cdot \mathbf{F}(\mathbf{r})$ of a vector field

$$\nabla \cdot \mathbf{F} \equiv \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \tag{4.20.1}$$

is a scalar field.

The curl $\nabla\times\mathbf{F}(\mathbf{r})$ of a vector field

$$\nabla\times\mathbf{F} \equiv \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}. \tag{4.20.2}$$

is another vector field.

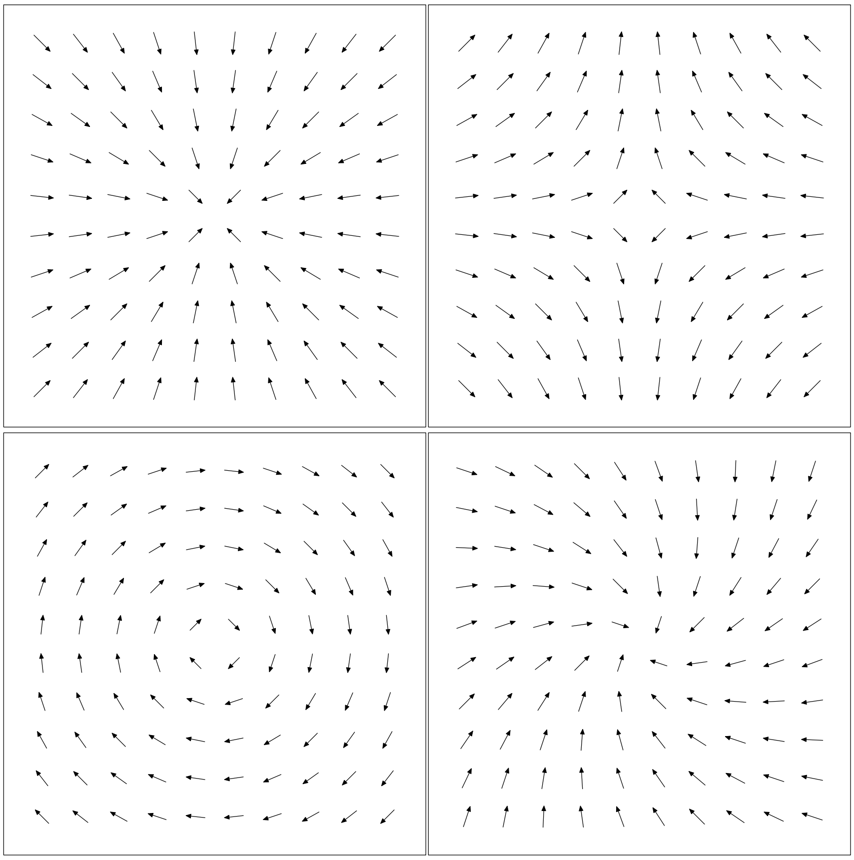

Figure 4.3: Illustration of some archetypical vector fields.

div $\neq 0$, curl $= 0$ “sink”	div $= 0$, curl $= 0$ “saddle”
div $= 0$, curl $\neq 0$ “cyclone”	div $\neq 0$, curl $\neq 0$ “whirlpool”

4.21 Interpretation of divergence and curl. Divergence and curl have useful geometrical interpretations, illustrated in Figure ??.

We will justify these interpretations in detail later.

4.22 Other derivatives. Other operations involving ∇ are also possible. In particular, the Laplacian

$$\nabla^2 \equiv \nabla \cdot \nabla \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{4.22.1}$$

and

$$\mathbf{F} \cdot \nabla \equiv F_x \frac{\partial}{\partial x} + F_y \frac{\partial}{\partial y} + F_z \frac{\partial}{\partial z} \tag{4.22.2}$$

we will meet again.

4.23 More identities. There are many useful identities involving gradient, divergence, and curl.

Curl of gradient is always zero, divergence of curl is always zero. These follow immediately on expanding components.

Another important identity is

$$\nabla \cdot (\Phi \mathbf{u}) = \Phi \nabla \cdot \mathbf{u} + \nabla \Phi \cdot \mathbf{u}. \tag{4.23.1}$$

Expanding out the components and using the shorthand

$$\partial_x \equiv \frac{\partial}{\partial x} \tag{4.23.2}$$

and so on, we have

$$\begin{aligned} &\partial_x(\Phi u_x) + \partial_y(\Phi u_y) + \partial_z(\Phi u_z) = \\ &\Phi(\partial_x u_x + \partial_y u_y + \partial_z u_z) + (u_x \partial_x \Phi + u_y \partial_y \Phi + u_z \partial_z \Phi) \end{aligned} \tag{4.23.3}$$

which is easily verified.

A related identity is

$$\nabla \times (\Phi \mathbf{u}) = \Phi \nabla \times \mathbf{u} + \nabla \Phi \times \mathbf{u} \tag{4.23.4}$$

Expanding out the components again we have

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ \Phi u_x & \Phi u_y & \Phi u_z \end{vmatrix} = \Phi \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ u_x & u_y & u_z \end{vmatrix} + \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x \Phi & \partial_y \Phi & \partial_z \Phi \\ u_x & u_y & u_z \end{vmatrix} \tag{4.23.5}$$

which is boring but not hard to verify.

4.24 Still more identities. Other vector identities are

$$\begin{aligned} &\nabla \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot \nabla \times \mathbf{u} - \mathbf{u} \cdot \nabla \times \mathbf{v} \\ &\nabla \times (\mathbf{u} \times \mathbf{v}) = \mathbf{u} \nabla \cdot \mathbf{v} - \mathbf{v} \nabla \cdot \mathbf{u} + \mathbf{v} \cdot \nabla \mathbf{u} - \mathbf{u} \cdot \nabla \mathbf{v} \\ &\nabla \times \nabla \times \mathbf{u} = \nabla \, \nabla \cdot \mathbf{u} - \nabla \cdot \nabla \mathbf{u} \\ &\nabla \, (\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \times \nabla \times \mathbf{v} + \mathbf{v} \times \nabla \times \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{u}. \end{aligned} \tag{4.24.1}$$

Although it is possible to derive these using component-expansion, it is very messy. We will develop a more concise technique for identities later.

****4.25 Rotational covariance.** A vector's magnitude and direction exist independently of any coordinate system, but individual components depend on the coordinate system.

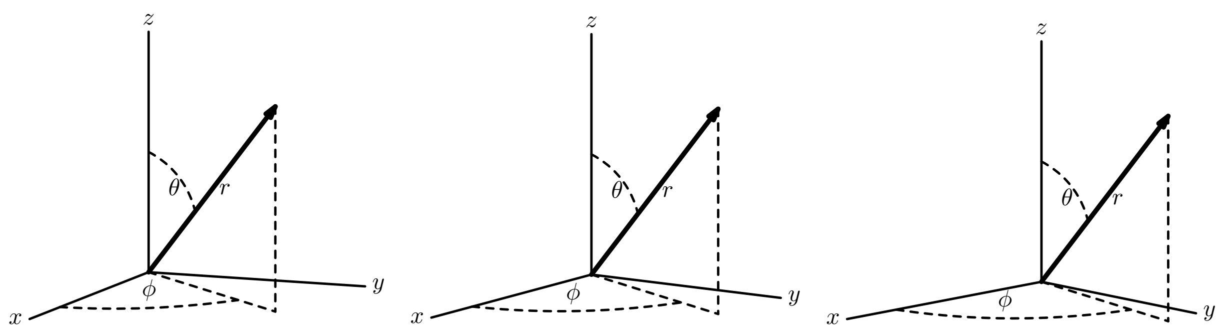

Figure 4.4: Coordinates may rotate (about z in this case) but vectors remain unchanged.

For example, rotating by ψ about the z axis (Figure 4.4) changes components as follows:

$$\begin{pmatrix} u'_x \\ u'_y \\ u'_z \end{pmatrix} = \begin{pmatrix} \cos\psi\, u_x + \sin\psi\, u_y \\ -\sin\psi\, u_x + \cos\psi\, u_y \\ u_z \end{pmatrix} \tag{4.25.1}$$

Dot products are invariant under rotation

$$\mathbf{u}' \cdot \mathbf{v}' = \mathbf{u} \cdot \mathbf{v} \tag{4.25.2}$$

while cross products transform like vectors:

$$\begin{pmatrix} (\mathbf{u}\times\mathbf{v})'_x \\ (\mathbf{u}\times\mathbf{v})'_y \\ (\mathbf{u}\times\mathbf{v})'_z \end{pmatrix} = \begin{pmatrix} \cos\psi\, (\mathbf{u}\times\mathbf{v})_x + \sin\psi\, (\mathbf{u}\times\mathbf{v})_y \\ -\sin\psi\, (\mathbf{u}\times\mathbf{v})_x + \cos\psi\, (\mathbf{u}\times\mathbf{v})_y \\ (\mathbf{u}\times\mathbf{v})_z \end{pmatrix} \tag{4.25.3}$$

***4.26 Example** [A simple rotation] Under rotation by 90° about z

$$x' = y,\ y' = -x,\ z' = z, \quad \mathbf{i}' = \mathbf{j},\ \mathbf{j}' = -\mathbf{i},\ \mathbf{k}' = \mathbf{k}. \tag{4.26.1}$$

Under this transformation $x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$ keeps the same form. (In fact, it does so under any rotation).

***4.27 Example** [An imposter] There is nothing to stop us defining a new kind of product

$$\mathbf{u} \circ \mathbf{v} \equiv u_x v_x\,\mathbf{i} + u_y v_y\,\mathbf{j} + u_z v_z\,\mathbf{k} \tag{4.27.1}$$

but nobody does.

The reason is that $\mathbf{u} \circ \mathbf{v}$ is not a vector, in that its magnitude and direction depend on coordinate system, making it useless. To see this, consider

$$\mathbf{r} \circ \mathbf{r} = x^2\,\mathbf{i} + y^2\,\mathbf{j} + z^2\,\mathbf{k}\ . \tag{4.27.2}$$

Under the coordinate rotation (4.26.1) it becomes

$$\mathbf{r}' \circ \mathbf{r}' = x'^2\,\mathbf{i}' + y'^2\,\mathbf{j}' + z'^2\,\mathbf{k}' = -x^2\,\mathbf{i} + y^2\,\mathbf{j} + z^2\,\mathbf{k} \tag{4.27.3}$$

which is inconsistent.

5. Index Notation

Index notation is a concise way of writing vectors. We develop it here for cartesian coordinates and components.

5.1 Coordinate indices. We introduce an alternative notation for coordinates and vectors. We write (x, y, z) as (x_1, x_2, x_3) or simply x_i, and (F_x, F_y, F_z) as (F_1, F_2, F_3) or simply F_i. In particular, x_i now denotes a position vector.

We can write a vector equation

$$\mathbf{w} = \mathbf{u} + \mathbf{v} \tag{5.1.1}$$

as

$$w_i = u_i + v_i \quad \text{or} \quad w_k = u_k + v_k. \tag{5.1.2}$$

An index used thus (called a free index) may change, but it must change consistently—$w_i = u_k + v_k$ is meaningless in index notation.

Indices that get summed over, as in

$$\sum_{i=1}^{3} A_i B_i + \sum_{i=1}^{3} C_i D_i \quad \text{or} \quad \sum_{j=1}^{3} A_j B_j + \sum_{k=1}^{3} C_k D_k. \tag{5.1.3}$$

are called dummy indices. These can change in a term independently of other terms.

5.2 Summation convention. In practice, dummy indices almost always appear in pairs within a term while free indices appear only once per term. It was Einstein, no less, who first noticed this fact and introduced the summation convention: an index may appear only once or twice per term; if once, it is understood to be a free index; if twice, it is understood to be a dummy index and a sum over 1 to 3 is implied.

It is very rare to need an index repeating in a term without being summed over, or an index appearing more than twice in a term. In such cases, we simply suspend the summation convention by writing *no sum* or similar next to the equation.

5.3 Example [Matrices] Index notation, including the summation convention, is useful not only for vectors. In particular we can easily prove the matrix identity $(AB)^T = B^T A^T$. Writing the matrix $C = AB$, we have

$$C^T_{ij} = C_{ji} = A_{jk} B_{ki} = A^T_{kj} B^T_{ik} = B^T_{ik} A^T_{kj}. \tag{5.3.1}$$

5.4 Dot products. A dot product $\mathbf{u} \cdot \mathbf{v}$ can be written as

$$u_i v_i = u_j v_j \tag{5.4.1}$$

5.5 Gradient, Divergence, and Laplacian. Let us define

$$\boxed{\partial_i \equiv \frac{\partial}{\partial x_i}} \tag{5.5.1}$$

Then the gradient of a scalar field ψ is

$$\partial_i \psi \tag{5.5.2}$$

while the divergence of a vector field $\mathbf{F}$ is

$$\partial_j F_j \tag{5.5.3}$$

and

$$\partial_j \partial_j \psi \tag{5.5.4}$$

denotes the Laplacian of ψ.

5.6 Example [Radius and its derivatives] In index notation $x_i x_i$ is just r^2. It follows that

$$\partial_i \, r^2 = 2x_i \tag{5.6.1}$$

and hence

$$\partial_i \, r^n = \partial_i \left((r^2)^{n/2} \right) = \tfrac{n}{2} (r^2)^{n/2-1} \, \partial_i \, r^2 = n r^{n-2} x_i \; . \tag{5.6.2}$$

We can further deduce the Laplacian

$$\partial_i \partial_i \, r^n = \partial_i \left(n r^{n-2} x_i \right) = n(n-2) r^{n-4} x_i x_i + 3n r^{n-2} = n(n+1) r^{n-2} \; . \tag{5.6.3}$$

5.7 Example [A vector identity revisited] If we write the vector identity

$$\nabla \cdot (\psi \mathbf{u}) = \psi \, \nabla \cdot \mathbf{u} + \nabla \psi \cdot \mathbf{u} \tag{5.7.1}$$

(cf. equation 4.23.1) in index notation we have

$$\partial_i (\psi u_i) = \psi \, \partial_i u_i + (\partial_i \psi) u_i \tag{5.7.2}$$

which is one-line derivation.

5.8 The Kronecker delta. This is defined as

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise.} \end{cases} \tag{5.8.1}$$

There are two very important identities involving the Kronecker delta:

$$\partial_i \, x_j = \delta_{ij} \tag{5.8.2}$$

and

$$\delta_{ij} A_j = A_i \tag{5.8.3}$$

and here A_j can be replaced by any expression with that index.

Another identity is

$$\delta_{ii} = 3. \tag{5.8.4}$$

5.9 The Levi-Civita symbol. This is ϵ_{ijk} defined as

$$\begin{cases} \epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1 \\ \epsilon_{321} = \epsilon_{132} = \epsilon_{213} = -1 \\ \text{all others} = 0. \end{cases} \tag{5.9.1}$$

In other words, $\epsilon_{123} = 1$ and ϵ_{ijk} is antisymmetric in any pair of indices.

A simple consequence is that for any S_{ij} symmetric in i, j

$$\epsilon_{ijk} S_{ij} = 0. \tag{5.9.2}$$

Also, because of antisymmetry $\epsilon_{iik} = 0$.

5.10 Cross product and curl. In index notation, a cross product becomes

$$(\mathbf{u}\times\mathbf{v})_i = \epsilon_{ijk} u_j v_k \tag{5.10.1}$$

and a curl becomes

$$(\nabla\times\mathbf{v})_i = \epsilon_{ijk} \partial_j v_k . \tag{5.10.2}$$

5.11 Example [Some vanishing identities] Using (5.9.2) is immediately follows that $\mathbf{u}\times\mathbf{u} = 0$ and also

$$(\nabla\times\nabla\psi)_i = \epsilon_{ijk} \partial_j \partial_k \psi = 0 \tag{5.11.1}$$

and

$$(\nabla \cdot \nabla\times\mathbf{F})_i = \partial_i \epsilon_{ijk} \partial_j F_k = 0. \tag{5.11.2}$$

5.12 Two Levi-Civita symbols. The Levi-Civita symbol satisfies a very important identity:

$$\boxed{\epsilon_{rmn} \epsilon_{rpq} = \delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}} \tag{5.12.1}$$

We prove the by verifying the possible cases.

(i) If all of $1, 2, 3$ are included in m, n, p, q then both sides give 0. The LHS gives 0 because r cannot be different from all of m, n, p, q. The RHS gives 0 because we cannot have $m = p, n = q$ and we cannot have $n = q, m = p$.

(ii) If $m = n$ or $p = q$ then both sides give 0. If only one of these is true, both terms on the RHS are 0; if both are true, the terms on the RHS cancel.

(iii) If $m = p \neq n = q$ then both sides give $+1$. The ϵ terms have the same sign.

(iv) If $m = q \neq n = p$ then both sides give -1.

5.13 Example [More identities] The identity (5.12.1) brings otherwise hard-to-prove identities quickly to heel. For instance the vector triple product (4.9.2) follows in two steps:

$$\begin{aligned} \epsilon_{ijk}\ \epsilon_{klm}\, A_j B_l C_m &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) A_j B_l C_m \\ &= A_j B_i C_j - A_j B_j C_i \ . \end{aligned} \tag{5.13.1}$$

Even the hardest of the standard identities

$$\nabla\,(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\times\nabla\times\mathbf{v} + \mathbf{v}\times\nabla\times\mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{u} \tag{5.13.2}$$

is not very difficult: we start with

$$\epsilon_{ijk}\, u_j\, \epsilon_{klm}\, \partial_l v_m = u_j \partial_i v_j - u_j \partial_j v_i \tag{5.13.3}$$

and then swap **u** and **v** and add.

6. Curved Coordinates

If a problem has a very non-rectangular geometry, it is often helpful to use non-cartesian coordinates, and cylindrical and spherical polar coordinates are particularly important examples of these.

6.1 Cylindrical coordinates. These are (ρ, ϕ, z) where

$$\boxed{x = \rho\cos\phi,\ \ y = \rho\sin\phi} \tag{6.1.1}$$

whose inverse is[3]

$$\rho = \sqrt{x^2 + y^2},\ \ \phi = \arctan(y, x). \tag{6.1.2}$$

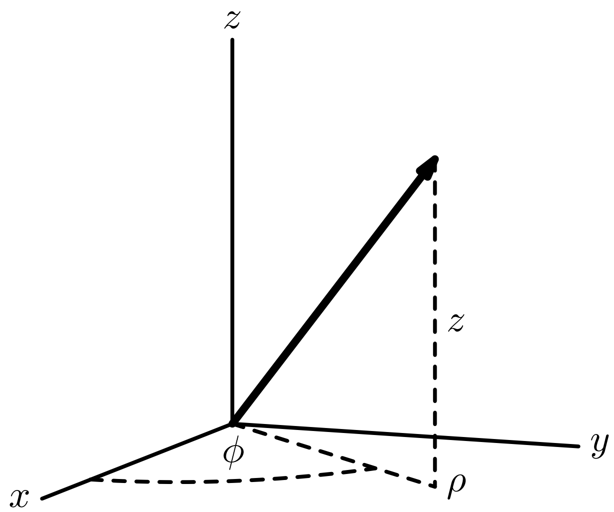

Figure 6.1: Cylindrical coordinates.

6.2 Spherical polar coordinates. These are (r, θ, ϕ) where

$$\boxed{x = r\sin\theta\cos\phi,\ \ y = r\sin\theta\sin\phi,\ \ z = r\cos\theta} \tag{6.2.1}$$

with inverse

$$r = \sqrt{x^2 + y^2 + z^2},\ \ \theta = \arctan\left(\frac{\sqrt{x^2 + y^2}}{z}\right),\ \ \phi = \arctan(y, x). \tag{6.2.2}$$

[3] Note the two-argument form of arctan here.

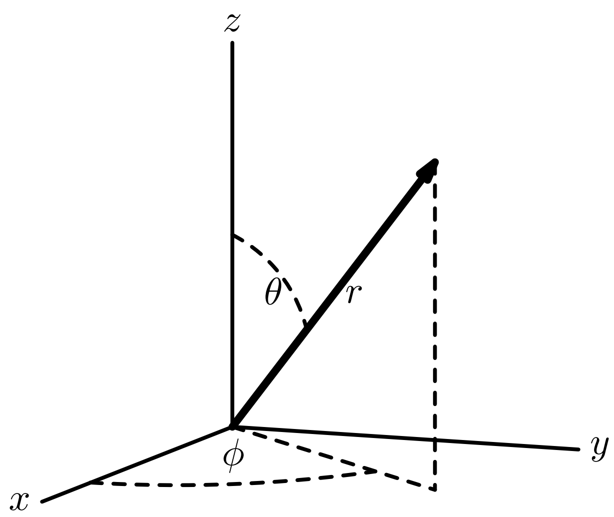

Figure 6.2: Spherical polar coordinates.

6.3 Example [Earth polar coordinates] Let the Earth's axis be the z axis with z increasing to the North, let the equator define the x, y plane, and let the prime meridian be $\phi = 0$. Then any point on the Earth can be referred to by the spherical polar angles (θ, ϕ), with ϕ being basically the longitude and θ being similar to the latitude but differing by $90°$.

For example Buenos Aires, which has latitude 35°S, and longitude 58°W, will have Earth polar coordinates $\theta = 125°, \phi = 302°$.

6.4 Coordinate curves and unit vectors. These are generalizations of coordinate axes: one of the coordinates varies while the others stay fixed. For example, in Earth polar coordinates, latitudes and longitudes are coordinate curves of θ and ϕ respectively.

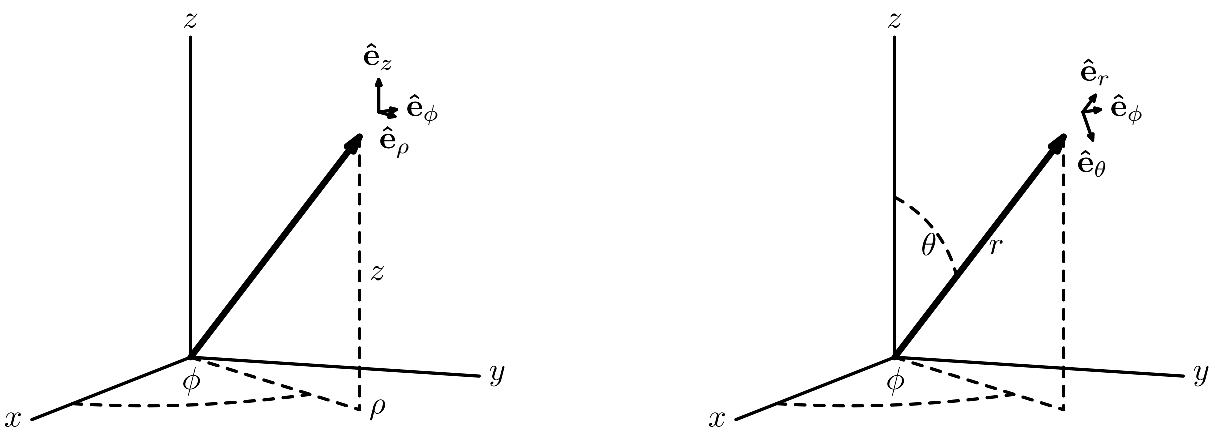

Figure 6.3: Unit vectors for cylindrical coordinates (left) and spherical polars (right).

Unit vectors tangent to the coordinate curves are defined as basis vectors analogous to **i**, **j**, **k**. We will derive expressions for them a little later, but for now see Figure 6.3 to get the idea.

We can expand any vector in terms of cylindrical or spherical polar coordinates using these unit vectors

$$\begin{aligned}\mathbf{F} &= F_x\,\mathbf{i} + F_y\,\mathbf{j} + F_z\,\mathbf{k} \\ &= F_\rho\,\hat{\mathbf{e}}_\rho + F_\phi\,\hat{\mathbf{e}}_\phi + F_z\,\hat{\mathbf{e}}_z \\ &= F_r\,\hat{\mathbf{e}}_r + F_\theta\,\hat{\mathbf{e}}_\theta + F_\phi\,\hat{\mathbf{e}}_\phi.\end{aligned} \tag{6.4.1}$$

Apart from $\hat{\mathbf{e}}_z$ (which is the same as $\mathbf{k}$) these new unit vectors all change direction depending on where you are.

***6.5 General unit vectors.** Instead of x, y, z we may have any three coordinates (say u_1, u_2, u_3) that uniquely specify a point in space. The unit vectors (or basis vectors), being tangents to the coordinate curves, are given by

$$\hat{\mathbf{e}}_i \equiv \frac{1}{h_i}\left(\frac{\partial \mathbf{r}}{\partial u_i}\right) \quad \text{where} \quad h_i \equiv \left|\frac{\partial \mathbf{r}}{\partial u_i}\right| \tag{6.5.1}$$

and the h_i are known as *scale factors*.

***6.6 Cylindrical unit vectors.** Substituting from (6.1.1) into the definition (6.5.1) and working through the algebra, we can derive

$$\begin{pmatrix}\hat{\mathbf{e}}_\rho \\ \hat{\mathbf{e}}_\phi \\ \hat{\mathbf{e}}_z\end{pmatrix} = \begin{pmatrix}\cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\mathbf{i} \\ \mathbf{j} \\ \mathbf{k}\end{pmatrix}. \tag{6.6.1}$$

As is easily verified, $(\hat{\mathbf{e}}_\rho, \hat{\mathbf{e}}_\phi, \hat{\mathbf{e}}_z)$ are orthogonal. This property is very important. In particular, it makes the inverse of (6.6.1) trivial.

***6.7 Spherical polar unit vectors.** We derive these by substituting from (6.1.1) into the definition (6.5.1) and working through more algebra, to get

$$\begin{pmatrix}\hat{\mathbf{e}}_r \\ \hat{\mathbf{e}}_\theta \\ \hat{\mathbf{e}}_\phi\end{pmatrix} = \begin{pmatrix}\sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0\end{pmatrix}\begin{pmatrix}\mathbf{i} \\ \mathbf{j} \\ \mathbf{k}\end{pmatrix} \tag{6.7.1}$$

which we can also write as

$$\begin{pmatrix}\hat{\mathbf{e}}_\theta \\ \hat{\mathbf{e}}_\phi \\ \hat{\mathbf{e}}_r\end{pmatrix} = \begin{pmatrix}\cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\mathbf{i} \\ \mathbf{j} \\ \mathbf{k}\end{pmatrix}. \tag{6.7.2}$$

Again, we can see that the unit vectors are orthogonal, and the inverse of (6.7.1) is also trivial.

***6.8 Example** [The position vector again] Using the formulas for coordinates and unit vectors (6.1.1 and 6.6.1 for cylindrical, 6.2.1 and 6.7.1 for spherical) we can convert vector components to and from cartesian.

In particular, substituting into (6.7.1) we can immediately verify that

$$r\,\hat{\mathbf{e}}_r = \mathbf{r}. \tag{6.8.1}$$

6.9 Dot and cross products. In any orthogonal coordinate system, these work just as in cartesian coordinates:

$$\mathbf{w}\cdot\mathbf{v} = w_1v_1 + w_2v_2 + w_3v_3 \tag{6.9.1}$$

and

$$\mathbf{w}\times\mathbf{v} = \begin{vmatrix} \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_3 \\ w_1 & w_3 & w_3 \\ v_1 & v_2 & v_3 \end{vmatrix}. \tag{6.9.2}$$

But differentiation is more complicated, because the unit vectors are no longer constant.

***6.10 Example** [Velocity in spherical polars] Using (6.8.1) we have

$$\dot{\mathbf{r}} = \dot{r}\,\hat{\mathbf{e}}_r + r\,\dot{\hat{\mathbf{e}}}_r \tag{6.10.1}$$

and using (6.7.1) to simplify the second term we get

$$\dot{\mathbf{r}} = \dot{r}\,\hat{\mathbf{e}}_r + r\dot{\theta}\,\hat{\mathbf{e}}_\theta + r\sin\theta\,\dot{\phi}\,\hat{\mathbf{e}}_\phi. \tag{6.10.2}$$

6.11 Scale factors. In sections 6.6 and 6.7 the scale factors defined in (6.5.1) were just normalization factors to be got rid of, but here they are:

$$\boxed{\begin{array}{lll} h_\rho = 1 & h_\phi = \rho & h_z = 1 \\ h_r = 1 & h_\theta = r & h_\phi = r\sin\theta \end{array}} \tag{6.11.1}$$

They are very useful, for the following reason. In any orthogonal coordinate system (u_1, u_2, u_3), small displacements along u_1 and u_2 define small reactangles, while small displacements along u_1, u_2, u_3 define small rectangular boxes. In other words, $h_1h_2\,du_1\,du_2$ is an area element and $h_1h_2h_3\,du_1\,du_2\,du_3$ is a volume element. The area and volume elements

$$\boxed{dx\,dy = \rho\,d\rho\,d\phi \qquad dx\,dy\,dz = r^2\sin\theta\,dr\,d\theta\,d\phi} \tag{6.11.2}$$

are especially useful.

6.12 Example [Circles and spheres] Using cylindrical coordinates the area within a circle of radius R comes out immediately as πR^2. Using spherical polar coordinates the volume of a sphere of radius R comes to $\frac{4}{3}\pi R^3$.

***6.13 Example** [Area of a cone] Consider the conical surface $\theta = \theta_1$ cut in a sphere of radius s. The area is given by integrating $(h_r\,dr)(h_\phi\,d\phi)$ or

$$\int_0^{2\pi} d\phi \int_0^s \sin\theta_1 r\,dr = \pi s^2 \sin\theta_1. \tag{6.13.1}$$

Here s is the slant height of a cone. The cone's base (say b) is going to be $s\sin\theta_1$. Hence we can express the sloping area of a cone neatly as $\pi s b$.

***6.14 Example** [Gaussian integral] Consider the integral

$$\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,dx\,dy = \left(\int_{-\infty}^{\infty} e^{-x^2}\,dx\right)^2. \tag{6.14.1}$$

Transforming to cylindrical coordinates gives

$$\int_0^\infty \rho e^{-\rho^2}\,d\rho \int_0^{2\pi} d\phi = \pi \tag{6.14.2}$$

and hence the most beautiful of all integrals

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}. \tag{6.14.3}$$

***6.15 Example** [Cutting an apple] Matthews poses an excellent 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 introducing $\sin\theta_1 = a/b$ and letting the radius of the apple equal unity. In our restated problem the corer touches the peel at

$$\rho = \sin\theta_1, \quad \theta = \theta_1, \quad z = \cos\theta_1. \tag{6.15.1}$$

We may also factor out 4π, by (i) by noting that integration over ϕ trivially gives 2π, and (ii) doing the integrals only for $z = 0$, and then doubling.

We now do this problem in four different ways.[4]

The first way is to integrate over z and then ρ

$$4\pi \int_0^{\sin\theta_1} \rho\,d\rho \int_0^{\sqrt{1-\rho^2}} dz = 4\pi \int_0^{\sin\theta_1} \rho(1-\rho^2)^{\frac{1}{2}}\,d\rho = \tfrac{4\pi}{3}(1-\cos^3\theta_1). \tag{6.15.2}$$

[4] The following gives only the key steps, quite a lot of algebraic filling-in is needed.

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

$$4\pi\int_{\cos\theta_1}^{1} dz\int_0^{\sqrt{1-z^2}}\rho\,d\rho = 2\pi\int_{\cos\theta_1}^{1}(1-z^2)\,dz = \tfrac{2\pi}{3}(2+\cos^3\theta_1-3\cos\theta_1). \quad (6.15.3)$$

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

$$4\pi\int_0^{\theta_1}\sin\theta\,d\theta\int_0^1 r^2\,dr = \tfrac{4\pi}{3}(1-\cos\theta_1). \quad (6.15.4)$$

Volume (ii), a cylinder with cone removed, is a bit harder:

$$4\pi\int_0^{\cos\theta_1} dz\int_{z\tan\theta_1}^{\sin\theta_1}\rho\,d\rho = 2\pi\int_0^{\cos\theta_1}(\sin^2\theta_1-z^2\tan^2\theta_1)\,dz = \tfrac{4\pi}{3}\sin^2\theta_1\cos\theta_1 \quad (6.15.5)$$

(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

$$4\pi\int_0^{\cos\theta_1} dz\int_{\sin\theta_1}^{\sqrt{1-z^2}}\rho\,d\rho = 2\pi\int_0^{\cos\theta_1}(1-z^2-\sin^2\theta_1)\,dz = \tfrac{4\pi}{3}\cos^3\theta_1. \quad (6.15.6)$$

7. Vector integration

We introduce line integrals, surface integrals, and volume integrals.

7.1 Line integrals. These are integrals along lines or curves, open or closed ($\int$ or $\oint$). The type

$$\int \mathbf{F}\cdot d\mathbf{r} \equiv \int \mathbf{F}\cdot \frac{d\mathbf{r}}{ds}\, ds \quad \text{along } \mathbf{r} = \mathbf{r}(s) \tag{7.1.1}$$

are the most important, and the type

$$\int \mathbf{F}\times d\mathbf{r} \tag{7.1.2}$$

comes up occasionally.

In dynamics, if $\mathbf{F}$ is a force, then $\mathbf{F}\cdot d\mathbf{r}$ is the work done by the force.

When facing a line integral, we begin by writing down $\mathbf{r}(s)$ and $(d\mathbf{r}/ds)$ and then substitute; this reduces the line integral to an ordinary integral in s.

7.2 Example [A simple line integral] Consider the integral $\int \mathbf{F}\cdot d\mathbf{r}$ with

$$\mathbf{F} = \begin{cases} \dfrac{\mathbf{r}}{r\sqrt{1-r^2}} \text{ if } r < 1 \\ 0 \text{ otherwise} \end{cases} \tag{7.2.1}$$

and taken along the positive z axis.

The path is the part of the z axis where the integrand is nonzero:

$$\mathbf{r} = t\,\mathbf{k}, \quad t \in [0,1] \tag{7.2.2}$$

and $d\mathbf{r} = \mathbf{k}\,dt$. Substituting gives

$$\int_0^1 \frac{dt}{\sqrt{1-t^2}} = \tfrac{\pi}{2}. \tag{7.2.3}$$

7.3 Example [Integral around a loop] Let us calculate

$$\oint \mathbf{r}\times d\mathbf{r} \tag{7.3.1}$$

as $\mathbf{r}$ goes around a loop of unit radius in the plane $z = c$.

Using cylindrical coordinates we have

$$\mathbf{r} = \cos\phi\,\mathbf{i} + \sin\phi\,\mathbf{j} + c\,\mathbf{k}, \quad 0 \le \phi \le 2\pi \tag{7.3.2}$$

and also

$$d\mathbf{r} = (-\sin\phi\,\mathbf{i} + \cos\phi\,\mathbf{j})\,d\phi. \tag{7.3.3}$$

The integral then reduces to

$$\oint \mathbf{r}\times d\mathbf{r} = \int_0^{2\pi} (-c\cos\phi\;\mathbf{i} - c\sin\phi\;\mathbf{j} + \;\mathbf{k})\,d\phi = 2\pi\,\mathbf{k}. \tag{7.3.4}$$

There is a slightly different thing we can do. For the given path, it is clear that $\mathbf{r}\times d\mathbf{r}$ is going to point along $\mathbf{k}$. So we may as well calculate the scalar integral

$$\mathbf{k}\cdot\oint \mathbf{r}\times d\mathbf{r} = \oint (\,\mathbf{k}\times\mathbf{r})\,d\mathbf{r}. \tag{7.3.5}$$

The last equality follows from the identity

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = (\mathbf{A}\times\mathbf{B})\cdot\mathbf{C} \tag{7.3.6}$$

which is easy to derive using index notation. Working out $(\,\mathbf{k}\times\mathbf{r})$ and simplifying leads to 2π for the integral (7.3.5).

7.4 Conservative vector fields. If a vector field can be expressed as a gradient then its line integral will depend only on the endpoints, not on the path. The converse also holds.

$$\boxed{\int \mathbf{F}\cdot d\mathbf{r}\ \text{path-independent} \Leftrightarrow \mathbf{F} = \nabla\Phi} \tag{7.4.1}$$

Vector fields with this property are called conservative, (because they often arise in dynamics problems, where they conserve energy).

We recall (from section 4.23) that curl of a gradient is zero. Hence

$$\nabla\times\mathbf{F} \neq 0 \Rightarrow \int \mathbf{F}\cdot d\mathbf{r}\ \text{path-}\textit{dependent.} \tag{7.4.2}$$

In fact the converse is also true, but we will come to that later.

Caution: These results apply only to vector fields that are well-behaved. If there is a discontinuity or other singularity then all bets are off.

****7.5 Path independence vs gradient.** Here we prove the equivalence (7.4.1).

To prove the '$\Leftarrow$' suppose $\mathbf{F} = \nabla\Phi$. Then by the chain rule

$$\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F}\cdot d\mathbf{r} = \int_{\mathbf{r}_1}^{\mathbf{r}_2} d\Phi = \Phi(\mathbf{r}_2) - \Phi(\mathbf{r}_1). \tag{7.5.1}$$

To prove the '$\Rightarrow$' define

$$\Phi(\mathbf{r}) = \int_{\mathbf{r}_1}^{\mathbf{r}} \mathbf{F}\cdot d\mathbf{r} = \int_{\mathbf{r}_1}^{\mathbf{r}} \mathbf{F}\cdot \frac{d\mathbf{r}}{dt}\; dt \tag{7.5.2}$$

along some path $\mathbf{r}(t)$. Differentiating gives

$$\frac{d\Phi}{dt} = \mathbf{F} \cdot \frac{d\mathbf{r}}{dt} \,. \tag{7.5.3}$$

Meanwhile

$$\frac{d\Phi}{dt} = \nabla\Phi \cdot \frac{d\mathbf{r}}{dt} \tag{7.5.4}$$

follows from the chain rule. Since the last two equations are valid for any path

$$\mathbf{F} = \nabla\Phi. \tag{7.5.5}$$

7.6 Example [Two paths] We consider two vector fields

$$\mathbf{E} = y\,\mathbf{i}, \quad \mathbf{F} = x\,\mathbf{j} \tag{7.6.1}$$

and integrate them from $(0,0)$ to $(1,1)$ along two different paths.

One path consists of

$$y = 0, \quad x = t, \quad t \in [0,1] \tag{7.6.2}$$

followed by

$$x = 1, \quad y = t, \quad t \in [0,1]. \tag{7.6.3}$$

This path gives

$$\int \mathbf{E} \cdot d\mathbf{r} = 0, \quad \int \mathbf{F} \cdot d\mathbf{r} = 1. \tag{7.6.4}$$

The other path consists of

$$x = y = t, \quad t \in [0,1] \tag{7.6.5}$$

and along this path

$$\int \mathbf{E} \cdot d\mathbf{r} = \int \mathbf{F} \cdot d\mathbf{r} = \tfrac{1}{2}. \tag{7.6.6}$$

Thus the line integrals of $\mathbf{E}$ and $\mathbf{F}$ depend on path.

On the other hand, notice that

$$\int (\mathbf{E} + \mathbf{F}) \cdot d\mathbf{r} = 1 \tag{7.6.7}$$

for both paths. This could have been predicted, because

$$\mathbf{E} + \mathbf{F} = \nabla(xy) \,. \tag{7.6.8}$$

***7.7 Example** [A loopy integral] Let us calculate $\int \mathbf{F} \cdot d\mathbf{r}$ for

$$\mathbf{F} = \frac{-y\,\mathbf{i} + x\,\mathbf{j}}{x^2 + y^2} \tag{7.7.1}$$

We try changing variables to

$$x = \rho \cos\phi, \ \ y = \rho \sin\phi \tag{7.7.2}$$

which gives

$$d\mathbf{r} = (\cos\phi\ \mathbf{i} + \sin\phi\ \mathbf{j})\, d\rho + \rho(-\sin\phi\ \mathbf{i} + \cos\phi\ \mathbf{j})\, d\phi + \ \mathbf{k}\, dz \tag{7.7.3}$$

and hence

$$\mathbf{F} \cdot d\mathbf{r} = d\phi. \tag{7.7.4}$$

The integral is $2n\phi$, where n is the number of times the path loops around the z axis.

Examining equation (7.7.1), we can verify that in fact $\mathbf{F} = \nabla\phi$. Yet we found its line integral depending on path, seemingly contradicting (7.4.1). The resolution is the restriction of (7.4.1) to well-behaved functions; ϕ is not well-behaved on the z-axis.

7.8 Surface integrals. The concept of a local normal to a surface (see Figure 4.2) makes possible a surface integral

$$\int \mathbf{F} \cdot d\mathbf{S}, \quad \mathbf{S} \equiv \hat{\mathbf{n}}\, dS \tag{7.8.1}$$

where $\hat{\mathbf{n}}$ is the unit normal and dS a surface-element. $\mathbf{F} \cdot \hat{\mathbf{n}}$ is sometimes called the flux.

If the flux is unity, we just get the surface area.

7.9 Example [Solid angle] Consider the vector field

$$\mathbf{F} = \frac{\hat{\mathbf{r}}}{r^2}\ . \tag{7.9.1}$$

Surface integrals of $\mathbf{F}$ represent solid angles subtended at the origin. In particular, if we take the surface integral over the sphere $x^2 + y^2 + z^2 = a^2$, we have

$$\oint \mathbf{F} \cdot d\mathbf{S} = \int \sin\theta\, d\theta\, d\phi \tag{7.9.2}$$

which gives 4π for the solid angle subtended by a sphere.

7.10 Volume integrals. These are integrals of the type

$$\int F(x,y,z)\ dx\, dy\, dz \quad \text{also written as} \quad \int F(\mathbf{r})\, d^3\mathbf{r} \quad \text{or} \quad \int F(\mathbf{r})\, dV. \tag{7.10.1}$$

There are no vector dot or cross products to worry about.

If $F(\mathbf{r}) = 1$ the integral is simply the volume.

7.11 Example [Another Gaussian integral] Consider the integral

$$\int e^{-r^2}\, d^3\mathbf{r} \tag{7.11.1}$$

over all space. We can reduce this to a Gaussian integral (section 6.14) in two ways. One way is to rewrite the integral as

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx \int_{-\infty}^{\infty} e^{-y^2}\, dy \int_{-\infty}^{\infty} e^{-z^2}\, dz = \pi^{3/2}. \tag{7.11.2}$$

Another way is to change to spherical polars. Then the θ and ϕ integrals are trivial, leaving us with

$$4\pi \int_0^\infty r^2 e^{-r^2}\, dr \tag{7.11.3}$$

Now we integrate by parts, noting that $\int re^{-r^2}\, dr = -\frac{1}{2}e^{-r^2}$:

$$\int_0^\infty r\, re^{-r^2}\, dr = r\left(-\tfrac{1}{2}e^{-r^2}\right)\Big|_0^\infty + \tfrac{1}{2}\int_0^\infty e^{-r^2}\, dr = \tfrac{1}{4}\sqrt{\pi} \tag{7.11.4}$$

using the Gaussian integral in the last step. Hence the original integral is $\pi^{3/2}$.

8. Integral Theorems

Two theorems relating line, surface, and volume integrals are central to vector calculus. They are Stokes' theorem and the Divergence theorem (also called Gauss's theorem).

8.1 Well-behaved functions. When we say $f(x, y, z)$ is well behaved at (x, y, z) we mean that it has a Taylor expansion

$$f(x + \Delta x, y + \Delta y, z + \Delta z) = f(x, y, z)+ \left[\frac{\partial}{\partial x} f(x, y, z)\right] \Delta x + \left[\frac{\partial}{\partial y} f(x, y, z)\right] \Delta y + \left[\frac{\partial}{\partial z} f(x, y, z)\right] \Delta z + \langle \text{higher order} \rangle . \tag{8.1.1}$$

(But other sources may mean other things.)

8.2 Stokes' theorem. This is a very important theorem relating line and surface integrals:

$$\boxed{\oint \mathbf{F} \cdot d\mathbf{r} = \int (\nabla \times \mathbf{F}) \cdot d\mathbf{S}} \tag{8.2.1}$$

Here the line integral is taken over the boundary of the surface. (Hence different surfaces sharing a common boundary will give the same surface integral, if the theorem applies.)

Below we will prove Stokes' theorem for well-behaved vector fields. A sharper form of the theorem requires only that $\mathbf{F}$ has continuous derivatives, but we won't prove that.

****8.3 Line integral around a small rectangle.** Consider the line integral $\oint \mathbf{F} \cdot d\mathbf{r}$ of a well-behaved vector field around a small rectangle with corners (x, y, z) and $(x + \Delta x, y + \Delta y, z)$, as in Figure ??.

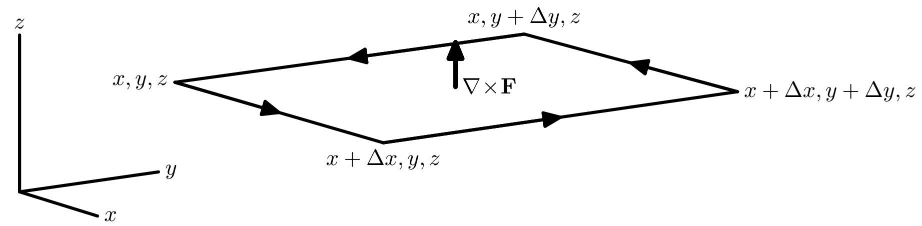

Figure 8.1: The relation between line and surface integrals.

The contribution from the two segments along x is

$$\left[F_x \Delta x + \left(\frac{\partial F_x}{\partial x}\right) \frac{\Delta x^2}{2}\right] - \left[F_x \Delta x + \left(\frac{\partial F_x}{\partial x}\right) \frac{\Delta x^2}{2} + \left(\frac{\partial F_x}{\partial y}\right) \Delta y \Delta x\right] = -\left(\frac{\partial F_x}{\partial y}\right) \Delta x \Delta y. \tag{8.3.1}$$

[We are neglecting higher orders and it is understood that F_x and so on are evaluated at (x, y, z).] Similarly, the other two segments contribute

$$\left(\frac{\partial F_y}{\partial x}\right) \Delta x \Delta y. \tag{8.3.2}$$

Hence the total line integral is

$$\left[\left(\frac{\partial F_y}{\partial x}\right) - \left(\frac{\partial F_x}{\partial y}\right)\right] \Delta x \Delta y = (\nabla \times \mathbf{F})_z \, \Delta x \Delta y. \tag{8.3.3}$$

****8.4 Stokes' theorem (proof).** The previous section establishes Stokes' theorem for small rectangles. But we can build a large surface out of small rectangles and the integrals along the internal boundaries will all cancel. Hence Stokes' theorem applies to any closed curve and any surface bounded by it.

This neat proof relies on the vector field $\mathbf{F}$ in (8.2.1) being well-behaved. Most books give a stronger proof that only needs $\mathbf{F}$ to have continuous partial derivatives, but that a proof is more difficult.

8.5 Conservative vector fields revisited. A corollary of Stokes' theorem is that if a well behaved vector field has zero curl, then its line integrals are path independent. Hence we can now extend section 7.4 to say:

$$\boxed{\int \mathbf{F} \cdot d\mathbf{r} \text{ path-independent} \Leftrightarrow \mathbf{F} = \nabla \Phi \Leftrightarrow \nabla \times \mathbf{F} = 0} \tag{8.5.1}$$

8.6 Example [or counterexample?] Stokes' theorem does *not* apply if the surface contains a singularity. To see this, consider

$$\mathbf{F} = \frac{-y\,\mathbf{i} + x\,\mathbf{j}}{x^2 + y^2}. \tag{8.6.1}$$

We can easily verify that $\nabla \times \mathbf{F} = 0$ (except on the z axis). Meanwhile, we have already seen (section 7.7) that $\oint \mathbf{F} \cdot d\mathbf{r} \neq 0$ if we go around the z axis.

8.7 Green's theorem in the plane. An immediate corollary of Stokes theorem is that for well-behaved functions $X(x, y)$ and $Y(x, y)$

$$\oint X\,dx + Y\,dy = \int\!\!\int \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}\right) dx\,dy\,. \tag{8.7.1}$$

It follows from the z component of Stokes' theorem.

8.8 Example [Area within a curve] A surprising expression for the area inside a closed curve C in a plane is

$$\frac{1}{2}\oint_C \mathbf{r}\times d\mathbf{r}\ . \tag{8.8.1}$$

We can assume without loss of generality that the plane of the curve is the x, y plane. The z component of the integral is

$$\oint_C (x\,dy - y\,dx) \tag{8.8.2}$$

which by Green's theorem in the plane equals

$$\int_C dx\,dy = \langle\text{area inside } C\rangle\ . \tag{8.8.3}$$

8.9 The Divergence theorem. A relative of Stokes' theorem relates surface and volume integrals:

$$\boxed{\oint \mathbf{F}\cdot d\mathbf{S} = \int (\nabla\cdot\mathbf{F})\,dV} \tag{8.9.1}$$

Here the surface integral is taken over the boundary of the surface.

The identity (8.9.1) is called the Divergence theorem. Other names are Gauss's theorem and Ostrogradsky's theorem. As with Stokes' theorem, we will prove he result for well-behaved vector fields but a sharper form (which we won't prove) requires only that $\mathbf{F}$ have continuous derivatives.

****8.10 Surface integral over a small rectangular box..** Consider the surface integral $\oint \mathbf{F}\cdot d\mathbf{S}$ of a well-behaved vector field around a rectangular box (aka parallelepiped) as in Figure ??.

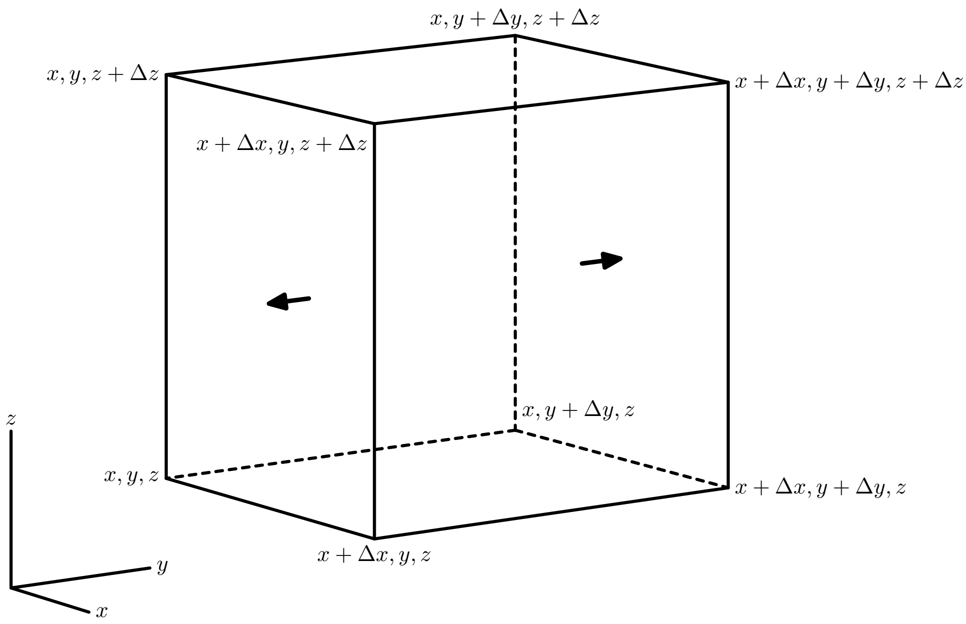

Figure 8.2: The relation between surface and volume integrals.

The net outward contribution along the two faces normal to y is

$$
\begin{aligned}
&- \left[F_y \Delta x \Delta z + \left(\frac{\partial F_y}{\partial x}\right) \frac{\Delta x^2}{2} \Delta z + \left(\frac{\partial F_y}{\partial z}\right) \frac{\Delta z^2}{2} \Delta x \right] \\
&+ \left[F_y \Delta x \Delta z + \left(\frac{\partial F_y}{\partial y}\right) \Delta y \Delta x \Delta z + \left(\frac{\partial F_y}{\partial x}\right) \frac{\Delta x^2}{2} \Delta z + \left(\frac{\partial F_y}{\partial z}\right) \frac{\Delta z^2}{2} \Delta x \right] \\
&= \frac{\partial F_y}{\partial y} \Delta x \Delta y \Delta z.
\end{aligned} \tag{8.10.1}
$$

Hence the total surface integral is

$$
\left[\left(\frac{\partial F_x}{\partial x}\right) + \left(\frac{\partial F_y}{\partial y}\right) + \left(\frac{\partial F_z}{\partial z}\right) \right] \Delta x \Delta y \Delta z = (\nabla \cdot \mathbf{F}) \, \Delta x \Delta y \Delta z. \tag{8.10.2}
$$

****8.11 The Divergence theorem (proof).** The previous section establishes the divergence theorem for small boxes. We can build a large volume out of small boxes and the integrals along the internal boundaries will all cancel. Hence the divergence applies to any volume and the surface bounding it.

As with the proof Stokes' theorem, this proof relies on the vector field $\mathbf{F}$ in (8.9.1) being well-behaved. And again, most books give a stronger but more difficult proof that only needs $\mathbf{F}$ to have continuous partial derivatives.

***8.12 Example** [Green's identities] We now derive two corollaries of the divergence theorem.

Let $f(\mathbf{r})$ and $g(\mathbf{r})$ be two well-behaved scalar fields. Then from the identity (4.23.1) we have

$$
\nabla \cdot (f \nabla g) = (\nabla f) \cdot (\nabla g) + f \nabla^2 g \tag{8.12.1}
$$

Taking (8.12.1) and its sibling identity with f and g swapped, and applying the divergence theorem gives

$$
\oint (f \nabla g - g \nabla f) \cdot d\mathbf{S} = \int (f \nabla^2 g - g \nabla^2 f) \, dV \tag{8.12.2}
$$

where integrals are over a closed surface and the volume enclosed by it.

For another identity, we put $g = f$ in (8.12.1) and apply the divergence theorem. We get

$$
\int (f \nabla f) \cdot d\mathbf{S} = \int \left((\nabla f)^2 + f \nabla^2 g \right) \, dV \tag{8.12.3}
$$

where again the integrals are over a closed surface and the volume enclosed by it. We are already given that f is well-behaved; let us further assume that as $r \to \infty$, f falls off faster than $1/\sqrt{r}$. Then $f \nabla f$ will fall off faster than $1/r^2$. Hence, if we take the integrals in (8.12.3) over a sphere of radius R, the LHS $\to 0$ as $R \to \infty$. Thus

$$
\int (\nabla f)^2 \, dV = - \int f \nabla^2 f \, dV \tag{8.12.4}
$$

where the integrals are over all space.

The identities (8.12.4) and (8.12.2) are known as Green's identities.

***8.13 Gauss's flux law.** This is an important corollary of the divergence theorem. It plays a key role in electromagnetism, where it relates an electric field to the charges that produce it.

Consider a vector field $\mathbf{F} = -\nabla(r^{-1}) = \hat{\mathbf{r}}/r^2$. As is easily verified, (i) $\mathbf{F}$ has zero divergence except at the origin, and (ii) over a sphere centred at the origin $\oint \mathbf{F}\cdot d\mathbf{S} = 4\pi$ (cf. section 7.9). It follows that

$$\mathbf{E} \equiv -\nabla\left(\frac{1}{|\mathbf{r}-\mathbf{r}_1|}\right) \qquad \oint \mathbf{E}\cdot d\mathbf{S} = \begin{cases} 4\pi & \text{if } \mathbf{r}_1 \text{ enclosed} \\ 0 & \text{otherwise .} \end{cases} \tag{8.13.1}$$

We can generalize to

$$\mathbf{E} \equiv -\nabla\Phi \qquad \Phi \equiv \sum_n \frac{q_n}{|\mathbf{r}-\mathbf{r}_n|} \qquad \oint \mathbf{E}\cdot d\mathbf{S} = 4\pi \sum_{n\ \text{encl}} q_n. \tag{8.13.2}$$

This is Gauss's flux law for discrete sources.

****8.14 Example** [A row of electric charges] Consider vector field

$$\mathbf{E} \equiv -\nabla\Phi \qquad \Phi \equiv \sum_{n=-\infty}^{\infty} \frac{1}{|\mathbf{r}-n\,\mathbf{k}|}. \tag{8.14.1}$$

We want to calculate $\mathbf{E}$ averaged along z:

$$\mathbf{E}_{(z\ \text{av})} \equiv \frac{\int \mathbf{E}\,dz}{\int dz}. \tag{8.14.2}$$

The average over z cannot depend on z. Also, the problem is symmetric ϕ. Hence $\mathbf{E}_{(z\ \text{av})}$ must have the direction of $\nabla(x^2+y^2)$, or $(x\,\mathbf{i}+y\,\mathbf{j})/\rho$. Using Gauss's flux law for a cylinder of radius ρ and length h, centred on the z axis and end faces avoiding the singularites, we have

$$2\pi\rho h\, E_{(z\ \text{av})} = 4\pi h \tag{8.14.3}$$

and hence

$$E_{(z\ \text{av})} = 2/\rho. \tag{8.14.4}$$

****8.15 Gauss's flux law for continuous sources.** It is possible to change from discrete sources q_n to a continuous source distribution $\rho(r)$, in which case (8.13.2) generalizes further to

$$\mathbf{E} \equiv -\nabla\Phi \qquad \Phi \equiv \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d^3\mathbf{r}' \qquad \oint \mathbf{E}\cdot d\mathbf{S} = 4\pi \int \rho(\mathbf{r})\, dV. \tag{8.15.1}$$

****8.16 Interpretation of divergence and curl.** We can rewrite (8.3.3) as

$$(\nabla\times\mathbf{F})_z = \frac{1}{\Delta x\Delta y}\left[\Delta x\,\frac{\partial}{\partial x}\,(F_y\Delta y) - \Delta y\,\frac{\partial}{\partial y}\,(F_x\Delta x)\right] \tag{8.16.1}$$

and $\Delta x\,(\partial F_y\Delta y/\partial x)$ expresses the change of the line integral along y and so on, while $\Delta x\Delta y$ represents area.

Analogously, we can express (8.10.2) as

$$\nabla\cdot\mathbf{F} = \frac{1}{\Delta x\Delta y\Delta z}\left[\Delta x\,\frac{\partial}{\partial x}\,(F_x\Delta y\Delta z) + \Delta y\,\frac{\partial}{\partial y}\,(F_y\Delta z\Delta x) + \Delta z\,\frac{\partial}{\partial z}\,(F_z\Delta x\Delta y)\right] \tag{8.16.2}$$

and here $\Delta x\,(\partial F_x\Delta y\Delta z/\partial x)$ and so on represent the change of outward flux, while $\Delta x\Delta y\Delta z$ represents volume.

8.17 Scale factors revisited. Scale factors tell us how derivatives and lengths transform to curved coordinates

$$\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \Delta x & \Delta y & \Delta z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \end{pmatrix} \longrightarrow \begin{pmatrix} \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_3 \\ h_1\Delta u_1 & h_2\Delta u_2 & h_3\Delta u_3 \\ \dfrac{1}{h_1}\dfrac{\partial}{\partial u_1} & \dfrac{1}{h_2}\dfrac{\partial}{\partial u_2} & \dfrac{1}{h_3}\dfrac{\partial}{\partial u_3} \end{pmatrix} \tag{8.17.1}$$

Combining with the interpretation of Stokes' theorem and the divergence theorem lets us express vector derivatives in any orthogonal coordinates.

8.18 Gradient in curved coordinates. Using (8.17.1) gives

$$\nabla\Phi = \hat{\mathbf{e}}_1\frac{1}{h_1}\frac{\partial\Phi}{\partial u_1} + \hat{\mathbf{e}}_2\frac{1}{h_2}\frac{\partial\Phi}{\partial u_2} + \hat{\mathbf{e}}_3\frac{1}{h_3}\frac{\partial\Phi}{\partial u_3}\,. \tag{8.18.1}$$

From (8.18.1) it follows that $\hat{\mathbf{e}}_1$ is always normal to a surface $f(u_2,u_3)=0$, and so on. So we could have defined the unit vectors as normals to coordinate surfaces rather than tangents to coordinate curves.

***8.19 Divergence in curved coordinates.** Using (8.17.1) and the interpretation of divergence (as in 8.16.2) lets us write

$$\nabla\cdot\mathbf{F} = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial u_1}\,(h_2h_3F_1) + \frac{\partial}{\partial u_2}\,(h_3h_1F_2) + \frac{\partial}{\partial u_3}\,(h_1h_2F_3)\right]. \tag{8.19.1}$$

***8.20 Example** [A unit-vector field] We can consider $\hat{\mathbf{e}}_r(\mathbf{r})$ as a vector field. It has divergence $2/r$, as we can verify by substituting $F_r = 1$ in (8.19.1).

***8.21 Curl in curved coordinates.** Using (8.17.1) and the interpretation of curl (as in 8.16.1) lets us write

$$(\nabla\times\mathbf{F})_3 = \frac{1}{h_1 h_2}\left[\frac{\partial}{\partial u_1}(h_2F_2) - \frac{\partial}{\partial u_2}(h_1F_1)\right] \tag{8.21.1}$$

and combining with analogous expressions for other components, we have

$$\nabla\times\mathbf{F} = \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\,\hat{\mathbf{e}}_1 & h_2\,\hat{\mathbf{e}}_2 & h_3\,\hat{\mathbf{e}}_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_2} \\ h_1F_1 & h_2F_2 & h_3F_3 \end{vmatrix}. \tag{8.21.2}$$

***8.22 Example** [A cylindrical identity] here we show that an arbitrary scalar field ψ satisfies

$$\nabla\cdot((\mathbf{k}\cdot\mathbf{r})(\mathbf{r}\times\nabla\psi)) = \frac{\partial\psi}{\partial\phi} \tag{8.22.1}$$

where $\mathbf{k}$ as usual is the unit vector along z.

To do this, we first work out

$$\begin{aligned}\nabla\cdot(\mathbf{r}\times\nabla\psi) &= \nabla\cdot\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ \psi_x & \psi_y & \psi_z \end{vmatrix} \\ &= y\psi_{zx} - z\psi_{yx} + z\psi_{xy} - x\psi_{zy} + x\psi_{yz} - y\psi_{xz} = 0.\end{aligned} \tag{8.22.2}$$

[Here subscripts denote partial derivatives.]

Next we note that

$$(\mathbf{k}\cdot\mathbf{r})(\mathbf{r}\times\nabla\psi) = z\,\mathbf{r}\times\nabla\psi\,. \tag{8.22.3}$$

Invoking the identity (4.23.1) we have

$$\nabla\cdot((\mathbf{k}\cdot\mathbf{r})(\mathbf{r}\times\nabla\psi)) = z\,\nabla\cdot(\mathbf{r}\times\nabla\psi) + \mathbf{k}\cdot\mathbf{r}\times\nabla\psi\,. \tag{8.22.4}$$

The first term is (as we just showed) zero. The second term is the z component of $\mathbf{r}\times\nabla\psi$. In cylindrical coordinates that means the z component of

$$\begin{vmatrix} \hat{\mathbf{e}}_\rho & \hat{\mathbf{e}}_\phi & \hat{\mathbf{e}}_z \\ \rho & 0 & z \\ (\partial\psi/\partial\rho) & (1/\rho)\,(\partial\psi/\partial\phi) & (\partial\psi/\partial z) \end{vmatrix}$$

or $(\partial\psi/\partial\phi)$.

***8.23 Laplacian in curved coordinates.** The Laplacian ∇^2 is divergence of gradient. Hence from (8.18.1) and (8.19.1) we have

$$\nabla^2 = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial u_1}\left(\frac{h_2h_3}{h_1}\frac{\partial}{\partial u_1}\right) + \frac{\partial}{\partial u_2}\left(\frac{h_3h_1}{h_2}\frac{\partial}{\partial u_2}\right) + \frac{\partial}{\partial u_3}\left(\frac{h_1h_2}{h_3}\frac{\partial}{\partial u_3}\right)\right] \tag{8.23.1}$$

which in cylindrical coordinates is

$$\nabla^2 = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2} \tag{8.23.2}$$

and in spherical polar coordinates is

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\,. \tag{8.23.3}$$

9. Laplace's equation

Laplace's equation is the simplest partial differential equation. It has many applications in physics and other areas.

9.1 The equation. Laplace's equation is

$$\boxed{\nabla^2\Phi(\mathbf{r}) = 0} \tag{9.1.1}$$

In general $\Phi(\mathbf{r})$ can depend on all three coordinates, but we will confine ourselves to cases depending on two variables: $\Phi(x, y)$ and $\Phi(\rho, \phi)$ and $\Phi(r, \theta)$.

Laplace's equation is a linear homogeneous equation. Hence a superposition of two solutions is also a solution.

9.2 Boundary conditions. There are infinite solutions of $\nabla^2\Phi = 0$, but the relevant ones are those for which Φ satisfies given boundary conditions. That is, we are given $\Phi(\mathbf{r})$ on a surface and we need to find $\Phi(\mathbf{r})$ either inside or outside that surface.

For $\Phi(x, y)$ or $\Phi(\rho, \phi)$ the boundary conditions are on walls of varying z at fixed x, y. But we can choose to forget the z direction and think of a planar problem, in which case there is a nice simple interpretation. Imagine Φ as a varying height. Then solving Laplace's equation subject to boundary conditions is like holding a rubber sheet between an arbitrarily shaped hoop. The hoop fixes the height at the boundary, while the rubber tries to minimize the total area.

For $\Phi(\rho, \phi)$ though, we still need to think in three dimensions, and the simple interpretation no longer applies.

9.3 Uniqueness theorem. For given boundary conditions, the solution of Laplace's equation is unique. Because of uniqueness, if we manage to guess something that satisfies both Laplace's equation and the boundary conditions, there are no other solutions to worry about.

***9.4 Uniqueness theorem (proof).** Suppose there are two solutions: $\nabla^2 f = 0$ and $\nabla^2 g = 0$. Define $\psi \equiv f - g$. Then (a) $\nabla^2\psi = 0$ and (b) $\psi = 0$ on the boundary. Now consider the identity

$$\nabla \cdot (\psi\nabla\psi) = (\nabla\psi)^2 + \psi\nabla^2\psi \,. \tag{9.4.1}$$

The last term is zero. Using the divergence theorem on the rest we have

$$\int (\psi\nabla\psi) \cdot d\mathbf{S} = \int (\nabla\psi)^2 \, dV \tag{9.4.2}$$

over the volume and boundary in question. Since $\psi = 0$ on the boundary, we infer successively that (i) the LHS is zero, (ii) $\nabla\psi = 0$ in the volume, and (iii) $\psi = 0$ in the volume. In other words $f = g$.

If the gradient rather than the value is specified on the boundary, the same proof works up to the conclusion (ii) above. The solution in this case is unique up to a constant.

9.5 Example [An electric shield] If a scalar field satisfies Laplace's equation inside a box and is constant on the surface of the box, then it must be constant inside the box.

If you stand under an electricity pylon, there is a rather voltage change—thousands of volts—between your head and your feet, and the voltage in space satisfies Laplace's equation. But if you stand inside a wire cage, then the wire has a property of equalizing the voltage over the cage and hence inside the cage too.

That is why the safest place from lightning is inside a wire cage.

9.6 Separating variables: cartesian coordinates. First of all, there is the almost trivial solution

$$\Phi_0(x, y) = \alpha + \beta x + \gamma y + \delta xy \tag{9.6.1}$$

with four constant coefficients. For a more interesting solution, suppose

$$\Phi(x, y) = X(x)Y(y) \; . \tag{9.6.2}$$

Substituting into (9.1.1) we get

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} = -\frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} = \text{const.} \tag{9.6.3}$$

Calling the constant $-k^2$ for definiteness we have

$$\Phi(x, y) = C_k \sin(kx + A_k) \sinh(ky + B_k) \; . \tag{9.6.4}$$

We can satisfy boundary conditions by superposing solutions of the type (9.6.1) and (9.6.4).

9.7 Example [A rectangular boundary] A typical problem is to solve

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\Phi = 0 \tag{9.7.1}$$

in a rectangular region where $\Phi(x, y)$ is specified on the boundary. The trick is to first subtract off a solution of the type (9.6.1) so that the corner values become zero, and deal with each nonzero side with contributions of the type (9.6.4).

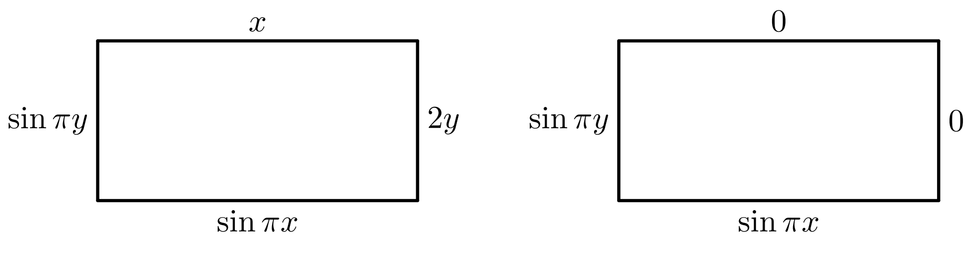

Figure 9.1: Left: boundary conditions on $\Phi(x, y)$. Right: boundary conditions after subtracting off $\Phi_0 = xy$.

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 Figure 9.1.

First we solve for the coefficients in $\Phi_0(x,y) = \alpha + \beta x + \gamma y + \delta xy$ to give $\Phi_0(0,0) = 0$, $\Phi_0(2,0) = 0$, $\Phi_0(0,1) = 0$, $\Phi_0(2,1) = 2$ and hence derive

$$\Phi_0 = xy \tag{9.7.2}$$

Substracting this off leaves the boundary conditions at the right of Figure 9.1. We can now get $\Phi(x,0)$ right with

$$\Phi_A(x,y) = \frac{\sinh(\pi(1-y))\sin \pi x}{\sinh(\pi)} \tag{9.7.3}$$

and get $\Phi(0,y)$ right with

$$\Phi_B(x,y) = \frac{\sinh(\pi(2-x))\sin \pi y}{\sinh(2\pi)} \ . \tag{9.7.4}$$

The full solution is

$$\Phi = \Phi_0 + \Phi_A + \Phi_B \ . \tag{9.7.5}$$

***9.8 Example** [Fourier series, again] Now consider a square with corners at $(0,0)$ and $(1,1)$ and boundary conditions as given in Figure (9.2).

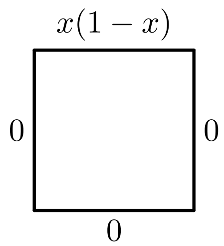

Figure 9.2: Boundary conditions.

This time the corner values are already zero. On the other hand $\Phi(x,1)$ is not conveniently a sinusoid. In this situation we need to expand the boundary condition in a Fourier series. Introducing a new variable $u = \pi x$, we have

$$\begin{aligned} x(1-x) &= \frac{u(\pi-u)}{\pi^2} = \sum_{n=1}^{\infty} b_n \sin nu \\ b_n &= \frac{2}{\pi^3}\int_0^{\pi} u(\pi-u)\,\sin nu\,du \ . \end{aligned} \tag{9.8.1}$$

Working out the integral, we have

$$x(1-x) = \sum_{n\ \mathrm{odd}}^{\infty} \frac{8}{n^3\pi^3}\sin(n\pi x) \tag{9.8.2}$$

and hence the solution is

$$\Phi(x,y) = \sum_{n\ \mathrm{odd}}^{\infty} \frac{8}{n^3\pi^3 \sinh(n\pi)} \sin(n\pi x)\sinh(n\pi y)\ . \tag{9.8.3}$$

***9.9 Cylindrical coordinates.** Now consider Laplace's equation for $\Phi(\rho,\phi)$ with no z dependence. Referring back to equation (8.23.2) we have

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\Phi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\Phi}{\partial\phi^2}\ . \tag{9.9.1}$$

Separating variables

$$\Phi(\rho,\phi) = F(\rho)\,G(\phi) \tag{9.9.2}$$

and substituting gives

$$\frac{\rho}{F}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial F}{\partial\rho}\right) = -\frac{1}{G}\frac{\partial^2 G}{\partial\phi^2} = \mathrm{const}\ m^2\ \text{(say)}. \tag{9.9.3}$$

For $m = 0$ we have

$$\Phi_0(\rho,\phi) = (A_0 + B_0\ln\rho)(C_0 + D_0\phi)\,. \tag{9.9.4}$$

If the origin is included and the solution is finite $B_0 = 0$, and if the solution is to be single-valued $D_0 = 0$. For nonzero m

$$\Phi_m(\rho,\phi) = (A_m\rho^m + B_m\rho^{-m})(C_m\cos m\phi + D_m\sin m\phi) \tag{9.9.5}$$

and for single-valued solutions m must be an integer.

The general solution is a superposition of solutions of the type (9.9.4) and (9.9.5).

***9.10 Example** [Around a cylinder] Consider $\nabla^2\Phi(\rho,\phi) = 0$ with the boundary condition $\Phi(1,\phi) = 2\sin^2\phi$ and $\Phi \sim \ln\rho$ at large ρ.

We have an $m = 0$ term and an $m = 2$ term, or

$$\Phi(\rho,\phi) = 1 + \ln\rho - \frac{\cos 2\phi}{\rho^2}\ . \tag{9.10.1}$$

***9.11 Spherical polar coordinates.** To finish with, we will consider Laplace's equation for $\Phi(r,\theta)$ with no ϕ dependence (axisymmetric solutions). Referring back to equation (8.23.3) we have

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) \tag{9.11.1}$$

Separating variables

$$\Phi(r,\theta) = F(r)\,G(\theta) \tag{9.11.2}$$

and substituting gives

$$\frac{1}{F}\frac{\partial}{\partial r}\left(r^2\frac{\partial F}{\partial r}\right) = -\frac{1}{G\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial G}{\partial\theta}\right) = \text{const } l(l+1) \text{ (say)}. \tag{9.11.3}$$

The radial part has solution

$$F(r) = A_l r^l + B_l r^{-(l+1)}\,. \tag{9.11.4}$$

For the angular part, we introduce a new variable

$$u = \cos\theta, \qquad \frac{d}{du} \equiv \frac{1}{\sin\theta}\frac{d}{d\theta} \tag{9.11.5}$$

and also write

$$P_l(u) = P_l(\cos\theta) = G(\theta) \tag{9.11.6}$$

to get

$$\frac{d}{du}\left((1-u^2)\frac{dP_l(u)}{du}\right) + l(l+1)P_l(u) = 0\,. \tag{9.11.7}$$

The general solution is a sum of terms

$$\Phi_l(r,\theta) = \left(A_l r^l + \frac{B_l}{r^{l+1}}\right)P_l(\cos\theta) \tag{9.11.8}$$

where $l \geq 0$.

***9.12 Legendre polynomials.** We still don't know what $P_l(u)$ is, except that it satisfies the differential equation (9.11.7). There are several ways of solving this equation. In particular, trying a series solution

$$\sum_{n=0}^{\infty} c_n u^n \tag{9.12.1}$$

gives

$$c_{n+2} = \frac{n(n+1)-l(l+1)}{(n+1)(n+2)}c_n\,. \tag{9.12.2}$$

Hence $P_l(u)$ is a polynomial of degree l. It is called a Legendre polynomial. The first few are

$$P_0(\cos\theta) = 1 \qquad P_1(\cos\theta) = \cos\theta \qquad P_2(\cos\theta) = \tfrac{1}{2}(3\cos^2\theta - 1) \tag{9.12.3}$$

****9.13 Example** [From electromagnetism] The physical situation of an earthed metal sphere in an external electric field gives rise to the following problem

$$\begin{aligned}&\Phi(r,\theta) = 0 \text{ for } r = a,\\ &\nabla^2\Phi(r,\theta) = 0 \text{ for } r > a,\\ &\Phi(r,\theta) \to r\cos\theta \text{ for large } r.\end{aligned} \tag{9.13.1}$$

The solution is

$$A\left(1-\frac{a}{r}\right)+\left(r-\frac{a^3}{r^2}\right)\cos\theta \tag{9.13.2}$$

Any further angular terms would destroy the constancy of Φ on the sphere. Note that A is still undetermined.

****9.14 Example** [Rodrigues formula] Legendre polynomials are important far beyond the context of Laplace's equation. An easier way of generating them is

$$P_l(u)=\frac{1}{2^l\,l!}\,\frac{d^l}{du^l}\,(u^2-1)^l \tag{9.14.1}$$

known as Rodrigues' formula.

With the help of Rodrigues' formula, it is not hard to verify that Legendre polynomials are orthogonal. They are the simplest of the so-called 'special functions' of mathematical physics.