1 Functions of several variables

Our course builds on the calculus that you have learnt in Single B. There you worked with functions of just one variable there; in this part of the course we will extend the idea of *differentiation and integration to functions of more than one variable.*

1.1 Examples

1. **Areas and volumes**;
   a) The volume V of a circular cylinder of radius x (cm) and height y (cm) is $V = \pi x^2 y$. A function of 2 variables. The cylinder gets larger if we increase x or y.
   b) A rectangular box with sides x, y, x has $V = xyz$ a function of 3 variables.

2. **Heights above a surface or displacements from a line**
   The position of a point on the earth given by ϕ, θ. (Latitude and longitude). The height above sea level can be expressed as $h(\theta, \phi)$.

3. **Atmospheric pressure**
   Again a function of θ, ϕ but also time dependent $P(t, \theta, \phi)$. Another variable? Variation of pressure with height r would be $P(t, r, \theta, \phi)$.

1.2 Graphs of functions

An equation $y = f(x)$ describes a curve in a plane. Given x we can compute y if we know the function f. Then we regard x and y as the coordinates of a point. The equation $y = f(x)$ describes how (x,y) moves as we vary x. The relation may be expressed implicitly as $g(x, y) = 0$.

1. $x^2 + y^2 = 1$ describes a circle of unit radius. This gives

$$y^2 = 1 - x^2$$

   so $y = \pm\sqrt{1 - x^2}$ for $|x| \leq 1$. Two equations of the form $y = f(x)$, one for the lower half one for the upper half.

2. In three dimensional space, $x^2 + y^2 + z^2 = 25$ describes a sphere of radius 5. This gives
$$z^2 = 25 - x^2 - y^2$$
so $z = \pm\sqrt{25 - x^2 - y^2}$ for $x^2 + y^2 \leq 25$. Note that for fixed y it describes circles in x, z plane. [How big are the circles if I slice 4 units from the origin? - $y = 4$ hence $z = \pm\sqrt{9 - x^2}$ and they are 3 units] Now have two equations of form $z = f(x, y)$. Given an x, y they give us two points lying on the sphere.

3. An equation of the form $z = f(x, y)$thus describes a surface. We can think of z as the height of the surface above the x, y plane. (or below if it is negative). The function f may be defined for all x, y or a restricted set as in the sphere. [What does $x^2 + y^2 > 25$ correspond to?]

1.3 Examples of graphs of functions

1. $z = 5$**:** for some range of x, y. - a flat roof. For the remaining examples see handouts

2. $z = x^2 + \sin y$: Defined for finite x and y. The lines show the curves on the surface corresponding to fixed x and y. For fixed y we get a parabola, and for fixed x we have a sin curve whose height is shifted.

3. $z = \cos(xy)$**:** Again the function is defined for all x, y. $z = const$ for $xy = const$. Contours of constant height would be hyperbolas. Keeping e.g. y constant gives cosine behaviour with a period determined by y.

4. $z = \frac{\sin\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}}$ the sombrero: Note the symmetry about the z axis through the origin. z has the same value for $x^2 + y^2 = r^2$ with r constant. This is the equation for a circle in the x, y plane.

5. $z = x^3 - 3xy^2$ the monkey saddle: Cubic curve for fixed y and parabolic for fixed x.

6. Three peaks: Combination of exponentials (Gaussians). It has maxima, minima and saddle points. We need to be able to classify them.

7. $z = (x^2 + y^2)/a^2$ satellite dish: This is a circular paraboloid. Parallel rays are focused onto the focus of the dish at $(0, 0, a)$.

2 Partial derivatives

For any function $f(x,y)$ we get

$\left(\frac{\partial f}{\partial x}\right)_y$ by differentiating w.r.t. x keeping y constant

$\left(\frac{\partial f}{\partial y}\right)_x$ by differentiating w.r.t. y keeping x constant

We often (in fact nearly always) drop the subscripts (when it is clear which variable is being held constant). Furthermore, sometimes $\frac{\partial f}{\partial x}$ is written f_x and $\frac{\partial f}{\partial y}$ is written f_y.

Differentiating again we get

$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$ also written f_{xx}

$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$ also written f_{yx}

$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$ also written f_{xy}

$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$ also written f_{yy}

For all "sensible" functions $f_{xy} = f_{yx}$.

2.1 Examples

1. $f(x,y) = x^3 - 3xy^2$. Then

$$\begin{aligned} f_x &= 3x^2 - 3y^2 \\ f_y &= -6xy \\ f_{xx} &= 6x \\ f_{yy} &= -6y \\ f_{xy} = f_{yx} &= -6y \end{aligned}$$

Note that $f_{xy} = f_{yx}$ for all reasonable functions and will be assumed from now on.

2. $f(x,y) = \cos y + \sin(xy)$. Then

$$\begin{aligned} f_x &= y\cos(xy) \\ f_y &= -\sin y + x\cos(xy) \\ f_{xx} &= -y^2\sin(xy) \\ f_{yy} &= -\cos y - x^2\sin(xy) \\ f_{xy} = f_{yx} &= \cos(xy) - xy\,\sin(xy) \end{aligned}$$

3. $f(x,y,z) = x^2y + yz + z^2x$. Then

$$\begin{aligned} f_x &= 2xy + z^2 \\ f_y &= z + x^2 \\ f_z &= y + 2xz \end{aligned}$$

2.2 Mathematical definition of the partial derivative

Recall the mathematical definition of the (standard) derivative:

$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

Mathematical definition of the partial derivative is similar:

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}$$

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y}$$

Recall that $\frac{df}{dx}$ gives the rate of change of f. Similarly $\frac{\partial f}{\partial x}$ corresponds to the rate of change of f **in the positive x direction**. And $\frac{\partial f}{\partial y}$ corresponds to the rate of change of f **in the positive y direction**.

(See handout for a picture of this.)

2.3 Multi-variable partial derivatives

Partial derivatives are defined for multi-variable functions in a similar way: For a function of n variables, $f(x_1, x_2, \ldots, x_n)$, we get:

$\frac{\partial f}{\partial x_i}$ by differentiating w.r.t. the variable x_i keeping **all other variables** constant

Mathematically we define

$$\frac{\partial f}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, x_2, \ldots, x_i + \Delta x_i, \ldots, x_n) - f(x_1, x_2, \ldots, x_i, \ldots, x_n)}{\Delta x_i}$$

2.4 The total differential

If $\frac{\partial f}{\partial x}$ corresponds to the rate of change of f **in the positive x direction** and $\frac{\partial f}{\partial y}$ corresponds to the rate of change of f **in the positive y direction**. It is natural to ask, What is the rate of change of f **in an arbitrary direction?**

Consider comparing the function at (x, y) , $f(x, y)$, with the function evaluated at a nearby point $(x + \Delta x, y + \Delta y)$ in order to obtain the rate of change of f in the arbitrary direction $(\Delta x, \Delta y)$. The change in f, Δf, is:

$$\begin{aligned}\Delta f &= f(x + \Delta x, y + \Delta y) - f(x, y) \\ &= \frac{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}{\Delta x}\Delta x + \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}\Delta y\end{aligned}$$

and as $\Delta x \to 0$ and $\Delta y \to 0$ this gets closer and closer to:

$$\Delta f \approx \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y$$

As $\Delta x \to 0$we can replace Δx with the infinitesimal dx, Δy with dx and Δf with df and write:

$$\boxed{df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy}$$

(The objects df, dx, dy are abstract mathematical objects. They are essentially represent the small changes $\Delta f, \Delta x, \Delta y$ when these become infinitesimally small. Such objects are known as **differentials**. It is important to get used to this concept even though it might seem strange at first sight! See Riley Ch2.1 for more info in the single variable case.)

We call df, defined by this equation, the **total differential** of f.

Example: what is the total differential of $f(x, y) = y \sin(x + y)$?
First compute the partial derivatives:

$$\frac{\partial f}{\partial x} = y\cos(x + y) \qquad\qquad \frac{\partial f}{\partial y} = y\cos(x + y) + \sin(x + y)$$

Then the total differential is

$$df = y\cos(x + y)\, dx + [y\cos(x + y) + \sin(x + y)]\ dy$$

Physical Example (Riley 5.11): The first law of thermodynamics can be expressed:

$$dE = T\,dS - P\,dV$$

where E is the internal energy of some substance, P its pressure, T temperature, V Volume and S Entropy. If we think of E, T, P as being functions of S, V relate $\left(\frac{\partial E}{\partial S}\right)_V$ and $\left(\frac{\partial E}{\partial V}\right)_S$ to T, P and hence prove the Maxwell relation:

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V .$$

Answer: Since we are thinking of E as a function of S, V we have that $dE = \left(\frac{\partial E}{\partial S}\right)_V \,dS + \left(\frac{\partial E}{\partial V}\right)_S \,dV$. Comparing with the first law gives immediately $T = \left(\frac{\partial E}{\partial S}\right)_V$ and $P = -\left(\frac{\partial E}{\partial V}\right)_S$. Now for the Maxwell relation, differentiate T wrt V to give $\left(\frac{\partial T}{\partial V}\right)_S = \frac{\partial^2 E}{\partial V \partial S}$ (drop subscripts here as it would get too cluttered, we must remember that S, V are our variables) on the other hand we get $\left(\frac{\partial P}{\partial S}\right)_V = -\frac{\partial^2 E}{\partial S \partial V}$ and since for second derivatives we have $\frac{\partial^2 E}{\partial S \partial V} = \frac{\partial^2 E}{\partial V \partial S}$ we immediately get the Maxwell relation. (Note: We will later get other Maxwell relations by considering different choices of our 2 variables.)

2.5 Exact and Inexact differentials

Not all differentials are total differentials of a function. For example $y\,dx + x\,dy$ is the total differential df of the function $f(x, y) = xy + c$ where c is a constant (**Check this!**). However try to find a function $f(x, y)$ whose total differential is $x\,dy + 3y\,dx$. It is impossible, there is no such function. Such a differential which is not the total differential of any function is known as an **inexact differential.** On the other hand a function which is the total differential of a function is known as an **exact differential.**

Example: show that $x\,dy + 3y\,dx$ is inexact.
For this to be exact it would have the form $f_x\,dx + f_y\,dy$ (here note that I am using the other notation for partial derivatives $f_x = \partial f/\partial x$ etc.) So $f_x = 3y$ and therefore (integrate wrt x) $f(x, y) = 3xy + g(y)$. On the other hand we must also have $f_y = x$ giving $f(x, y) = xy + h(x)$. These two results for $f(x, y)$ can not be made compatible and so no such $f(x, y)$ exists. The differential is thus inexact.

A useful criterion for exactness is the following:

A differential $A(x,y)\,dx + B(x,y)\,dy$ is exact if and only if $A_y = B_x$.

Note that one way around this statement is obvious, namely if $A(x,y)\,dx+B(x,y)\,dy$is exact (ie $= df$) then we have $A = f_x$and $B = f_y$and so $A_y = f_{yx} = f_{xy} = B_x$. The other way around we will use without proof. So we can now give a different proof of the above example:

Example: show that $x\,dy + 3y\,dx$ is inexact.
Here we have $A = 3y$ and $B = x$ so $A_y = 3$ and $B_x = 1$. These are not equal so the differential is thus inexact.

Example: Prove the Maxwell relation in the example above directly using this criterion and the fact that $dE = T\,dS - P\,dV$ is an exact differential.
Here we have variables S, T instead of x, y. Using the above criterion for an exact differential we have $A = T$ and $B = -P$ so the criterion $A_y = B_x$ becomes $T_V = -P_S$ which is the required Maxwell relation

2.5.1 The multi-variable case

This works similarly. The **total differential** of a multi-variable function $f(x_1, x_2, \ldots, x_n)$is

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i.$$

A differential

$$\sum_{i=1}^{n} g_i(x_1, x_2, \ldots, x_n) dx_i$$

is **exact** if it is the **total differential** of a function, ie if $g_i = \partial f / \partial x_i$ for some function f, and inexact otherwise. A necessary and sufficient condition for a differential to be exact is:

$$\frac{\partial g_i}{\partial x_j} = \frac{\partial g_j}{\partial x_i}.$$

For all $i, j = 1, \ldots, n$.

Example: show that $y(y+z)\,dx + x(2y+z)\,dy + xy\,dz$ is an exact differential. Here instead of $(x_1, x_2, x_3)$the variables are called (x, y, z) and we have $g_1 = y(y+z)$, $g_2 = 2xy + z$ and $g_3 = xy$. So $\frac{\partial g_1}{\partial x_2} = \frac{\partial g_1}{\partial y} = 2y + z = \frac{\partial g_2}{\partial x} = \frac{\partial g_2}{\partial x_1}$, similarly $\frac{\partial g_1}{\partial x_3} = y = \frac{\partial g_3}{\partial x_1}$, $\frac{\partial g_2}{\partial x_3} = x = \frac{\partial g_3}{\partial x_2}$. So $\frac{\partial g_i}{\partial x_j} = \frac{\partial g_j}{\partial x_i}$ for all i, j , so the differential is thus exact. You may be able to guess the function whose total differential it is. In this case $f = xy(y+z) + c$ for a constant c.

2.6 The chain rule

Usual case: If y is a function of a variable x and x is a function of another variable t then

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

Multi variable case: If $f = f(x, y)$ and x and y and both function of just t then in a time change Δt the change in e.g. x is $\Delta x = \frac{dx}{dt}\Delta t$ by definition. Change in f is then

$$\Delta f \approx \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y$$

Dividing through by Δt and taking the limit gives:

$$\boxed{\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}}$$

(Note that you can get this straight from the equation for the total differential by dividing by the infinitesimal dt.)

2.6.1 Examples

1. Suppose that the cylinder changes its size as $x = 3t$ and $y = 4 + t^2$. What is the rate of change of V? Method 1: Substitute:

$$\begin{aligned} V &= \pi x^2 y \\ &= \pi 9t^2(4 + t^2). \end{aligned}$$

   Then

$$\frac{dV}{dt} = 72\pi t + 36\pi t^3$$

Method 2: use chain rule

$$\begin{aligned}\frac{dV}{dt} &= \frac{\partial V}{\partial x}\frac{dx}{dt} + \frac{\partial V}{\partial y}\frac{dy}{dt} \\ &= 2\pi xy(3) + \pi x^2(2t) \\ &= \pi 18t(4+t^2) + \pi 18t^3 \\ &= 72\pi t + 36\pi t^3\end{aligned}$$

2. For $f = \sin(xy)$ find $\frac{df}{dt}$ along the curve parameterized by $x = t^2$, $y = t^3$.

$$\begin{aligned}\frac{df}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= y\cos(xy)(2t) + x\cos(xy)(3t^2) \\ &= 5t^4\cos(t^5)\end{aligned}$$

Note could have again substituted and used $f(t) = \sin(t^5)$

3. For $f(x, y, z) = 3xe^{yz}$ find the value of $\frac{df}{dt}$ at the point on the curve $x = \sin t$, $y = \cos t$, $z = t$ where $t = 0$.

$$\begin{aligned}\frac{df}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} \\ &= e^{yz}(3\cos t - 3xz\sin t + 3xz) \\ \frac{df}{dt}|_{t=0} &= 3.\end{aligned}$$

since at $t = 0$, $(x, y, z) = (0, 1, 0)$.

4. Let $f = f(x, t)$ where $x = x(t)$. What is $\frac{df}{dt}$?
The chain rule tells us that for functions $f(x, y)$ we have

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

For this example we can take $y(t) = t$. Then we have

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial t}$$

since $dt/dt = 1$. Note the importance of d/dt and $\partial/\partial t$

2.7 Change of variables

Suppose that $x = x(u, v)$ and $y = y(u, v)$ where u, v are two other variables. Then we have:

$$\begin{aligned}
\frac{\partial f}{\partial u} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \\
\frac{\partial f}{\partial v} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}
\end{aligned}$$

Note that $\frac{\partial f}{\partial x}$ *is at constant* y *but* $\frac{\partial x}{\partial u}$ *is at constant* v.

2.7.1 Examples

1. Let $f(x, y) = xy$ where $x = u \cos v$ and $y = u \sin v$. Substituting gives $f(u, v) = u^2 \sin v \, \cos v$ and $f_u = 2u \sin v \cos v$. The formula gives

$$\begin{aligned}
f_u &= f_x x_u + f_y y_u \\
&= y \cos v + x \sin v \\
&= 2u \sin v \cos v
\end{aligned}$$

   as required. It works.

2. If f is a function of x, y where $x = u^2 - v^2$ and $y = 2uv$ show that $u f_u - v f_v = 2(u^2 + v^2) f_x$. Use

$$\begin{aligned}
\frac{\partial f}{\partial u} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \\
&= 2u f_x + 2v f_y
\end{aligned}$$

   and

$$\begin{aligned}
\frac{\partial f}{\partial v} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} \\
&= -2v f_x + 2u f_y
\end{aligned}$$

   So that

$$\begin{aligned}
u f_u - v f_v &= 2u^2 f_x + 2uv f_y + 2v^2 f_x - 2uv f_y \\
&= 2(u^2 + v^2)\frac{\partial f}{\partial x}
\end{aligned}$$

   Note that since we do not know the form f a substitution is not even possible.

2.8 Some useful theorems of partial differentiation

In general, all variables are as good as any other. For functions of one variable we sometimes think of the inverse function. So if $y = f(x)$ then for the inverse, $x = f^{-1}(y)$. In other words there is no real difference in x and y, both are variables and we can treat them on equal footing as variables. We have been so far thinking mostly of a function f of x and y. We will now write z instead of f and treat x, y, z on an equal footing. So thinking of z as a function of x and y, $z(x, y)$ then we have:

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy$$

but we could equally invert the equation and write $x$as a function of y and z, $x(y, z)$

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz$$

or write y as a function of x and z, $y(x, z)$ and thus we have

$$dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz.$$

Plugging the last equation in to the previous one gives

$$dx = \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y\right) dz. \qquad *$$

Now this equation implies

$$\boxed{\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z = 1 \qquad \Rightarrow \qquad \left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial y}{\partial x}\right)_z^{-1}}$$

This is the analogue of the relation for single variable functions: $\frac{dy}{dx} = \left(\frac{dx}{dy}\right)^{-1}$, and is true for partial derivatives if the same variable(s) are kept fixed on both sides of the equation (in this case z).

The equation $*$ also implies the vanishing of the second term:

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y = 0 \qquad \Rightarrow \qquad \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x = -\left(\frac{\partial x}{\partial z}\right)_y = -\left(\frac{\partial z}{\partial x}\right)_y^{-1}$$

thus giving

$$\boxed{\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1}$$

Example: Show this identity for the equation $z = 2x + 3y$. Here $\left(\frac{\partial z}{\partial x}\right)_y = 2$, rearranging $x = \frac{1}{2}(z - 3y)$ so $\left(\frac{\partial x}{\partial y}\right)_z = -\frac{3}{2}$ and since $y = \frac{1}{3}(z - 2y)$, we have $\left(\frac{\partial y}{\partial z}\right)_x = \frac{1}{3}$. So $\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = (-\frac{3}{2})(\frac{1}{3})(2) = -1$.

2.9 Thermodynamic Example

We look at an example here where different quantities can become our variables. Previously we saw that the first law of thermodynamics takes the form

$$dE = T\,dS - P\,dV \tag{1}$$

where E is the internal energy of some substance, P its pressure, T temperature, V Volume and S Entropy. We then thought of E, T, P as being functions of S, V so S, V were our variables, and we then proved the Maxwell relation:

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V .$$

However it is possible to take any two quantities as our variables (depending on the physical situation, what are we physically varying in the experiment. Let us thus consider taking T, V as our variables, can we derive a different Maxwell relation?

Let us begin by writing all the differential forms in 1 in terms of these variables, so $dE = \left(\frac{\partial E}{\partial T}\right)_V dT + \left(\frac{\partial E}{\partial V}\right)_T dV$ and $dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV$ so we get

$$\left(\frac{\partial E}{\partial T}\right)_V dT + \left(\frac{\partial E}{\partial V}\right)_T dV = T\left(\left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV\right) - P\,dV$$

Now equate coeffs of dT and dV we get two equations: $\left(\frac{\partial E}{\partial T}\right)_V = T\left(\frac{\partial S}{\partial T}\right)_V$ and $\left(\frac{\partial E}{\partial V}\right)_T = T\left(\frac{\partial S}{\partial V}\right)_T - P$. Now differentiating the first equation wrt V and the second equation wrt T we get $\frac{\partial^2 E}{\partial V \partial T} = T\frac{\partial^2 S}{\partial V \partial T} = \left(\frac{\partial S}{\partial V}\right)_T + T\frac{\partial^2 S}{\partial V \partial T} - \left(\frac{\partial P}{\partial T}\right)_V$ giving immediately the Maxwell equation $\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V$. (We twice used the fact that for mixed second derivatives the order doesn't matter.)

3 Partial differential equations

An equation involving *functions of two or more variables and some of its spatial derivatives* is a p.d.e. Here we solve some simple examples that can be done easily.

3.1 Simple Examples

Example 1: $\frac{\partial u}{\partial x} = 0 \forall x$

The solution of the o.d.e. is $u = c$ a constant. But if $u = u(x, y)$ then integration with respect to x can give any function of y (check by reversing, i.e. differentiation). So then

$$u(x, y) = f(y)$$

where f is some function to be determined by the boundary conditions.,

Example 2a: Find the general sol'n of $\frac{\partial^2 u}{\partial x^2} = 0$ **where** $u = u(x, t)$**.**

Integration with respect to x gives $u_x = f(t)$ and again gives $u(x, t) = xf(t) + g(t)$, with f and g being arbitrary functions.

Example 2b: Find the sol'n of $\frac{\partial^2 u}{\partial x^2} = 0$ **when** $u(0, t) = 1$**,** $u(1, t) = \sin t$**.**

Must now determine f and g to satisfy *the boundary conditions.* Note that there are two "$t-$functions of integration" so two boundary conditions are required involving functions of t. We have

$$\begin{aligned} g(t) &= 1 \\ f(t) + g(t) &= \sin t \end{aligned}$$

so that $g(t) = 1$ and $f(t) = \sin t - 1$. The full solution is

$$u(x, t) = x(\sin t - 1) + 1.$$

Check this!

Example 3: Find the general solution $z(x,y)$ of $\frac{\partial^2 z}{\partial x \partial y} = 0$.

Integration w.r.t. x gives $z_y = F(y)$ an arbitrary function. And w.r.t. y gives

$$\begin{aligned} z &= \int^y F(y')dy' + g(x) \\ &= f(y) + g(x) \end{aligned}$$

where $f_y = F(y)$ and f and g are arbitrary *differentiable functions* (since z_{xy} must exist).

Example 4: Find the general solution $z(x,y)$ of $\frac{\partial^2 z}{\partial x \partial y} = x - y$.

Integration w.r.t x keeping y constant gives $z_y = \frac{x^2}{2} - xy + G(y)$. And with respect to y gives

$$z = \frac{xy}{2}(x-y) + f(x) + g(y)$$

where f and g are again arbitrary differentiable functions.

Example 5: Find the general solution $u(x,y)$ of $\frac{\partial u}{\partial x} + u = 2$.

How would we do this if it were an o.d.e? We multiply by integrating factor (since without the 2 the sol'n is just e^{-x}). Then

$$\begin{aligned} e^x u_x + e^x u &= 2e^x \\ \frac{\partial}{\partial x}(e^x u) &= 2\frac{\partial}{\partial x}(e^x) \\ ue^x &= 2e^x + f(y) \end{aligned}$$

so that

$$u = 2 + e^{-x} f(y).$$

Note you could also rearrange to get

$$\int \frac{du}{2-u} = \int dx$$

and remember to integrate keeping y constant

$$\begin{aligned} \log(2-u) &= c(y) - x \\ u &= 2 + e^{c(y)-x} \end{aligned}$$

with $f(y) \equiv e^{c(y)}$.

Example 6: Find the general solution $u(x, y)$ of $u_{xx} + \frac{u_x}{x} = 3x + 4$.

Let $p(x, y) = u_x$ then $xp_x + p = 3x^2 + 4x$ can be integrated straightaway

$$\begin{aligned} xp &= x^3 + 2x^2 + f(y) \\ u_x &= x^2 + 2x + \frac{f(y)}{x} \end{aligned}$$

and then integrating again gives

$$u = \frac{x^3}{3} + x^2 + f(y)\log(x) + g(y).$$

3.2 Aside: differentiation as an operator

It is extremely useful to view differentiation d/dt as an **operator** acting on the set of all functions and giving a new function called df/dt

$$\frac{d}{dt} : f \longmapsto \frac{df}{dt}$$

Then acting with d/dt on df/dt gives the second derivative

$$\frac{d}{dt} : \frac{df}{dt} \longmapsto \frac{d^2 f}{dt^2}.$$

The same is true for partial derivatives.

Always remember that derivatives act on everything to the right. For example, consider expanding out something like

$$\left(x^2\frac{\partial}{\partial x} + y^2\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right) f = x^2\frac{\partial^2 f}{\partial x^2} + y^2\frac{\partial^2 f}{\partial x \partial y} + x^2 y\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial}{\partial y}\left(y\frac{\partial}{\partial y}\right) f$$

where the last term gives

$$y^2\frac{\partial}{\partial y}\left(y\frac{\partial f}{\partial y}\right) = y^2\frac{\partial f}{\partial y} + y^3\frac{\partial^2 f}{\partial y^2}.$$

This means we can rewrite the change of variables rule as an operator equation:

$$\begin{aligned}\frac{\partial}{\partial u} &= \frac{\partial x}{\partial u}\frac{\partial}{\partial x} + \frac{\partial y}{\partial u}\frac{\partial}{\partial y}\\ \frac{\partial}{\partial v} &= \frac{\partial x}{\partial v}\frac{\partial}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial}{\partial y}\end{aligned}$$

Notice there are no f's in this equation. This can act on any function. Indeed it can act on df/dt which in turn can be obtained by acting with this operator on f. In other words we, for example, the following:

$$\frac{\partial^2 f}{\partial u^2} = \left(\frac{\partial x}{\partial u}\frac{\partial}{\partial x} + \frac{\partial y}{\partial u}\frac{\partial}{\partial y}\right)\left(\frac{\partial x}{\partial u}\frac{\partial}{\partial x} + \frac{\partial y}{\partial u}\frac{\partial}{\partial y}\right) f$$

where each operator acts on everything to the right of it.

Example: If $f(x, y)$ and $x(r, \theta) = r\cos\theta y(r, \theta) = r\sin\theta$ (polar coordinates) compute $\partial^2 f/\partial r^2$.

Using the formula for co-ordinate transformations above (with $u = r$ and $v = \theta$) we have:

$$\begin{aligned}\frac{\partial}{\partial r} &= \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y}\\ &= \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}\end{aligned}$$

So we have that

$$\begin{aligned}\frac{\partial}{\partial r}\left(\frac{\partial}{\partial r} f\right) &= \frac{\partial}{\partial r}\left(\cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}\right) f\\ &= \cos\theta\frac{\partial}{\partial r}\left(\frac{\partial f}{\partial x}\right) + \sin\theta\frac{\partial}{\partial r}\left(\frac{\partial f}{\partial y}\right)\end{aligned}$$

where we used the fact that $\frac{\partial\theta}{\partial r} = 0$ to move the $\frac{\partial}{\partial r}$ through the $\cos\theta$ and $\sin\theta$. Now we again use the above formula for $\frac{\partial}{\partial r}$ to give

$$\begin{aligned}\frac{\partial^2 f}{\partial r^2} &= \cos\theta\left(\cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}\right)\left(\frac{\partial f}{\partial x}\right) + \sin\theta\left(\cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}\right)\left(\frac{\partial f}{\partial y}\right)\\ &= \cos^2\theta\frac{\partial^2 f}{\partial x^2} + 2\cos\theta\sin\theta\frac{\partial^2 f}{\partial x\partial y} + \sin^2\theta\frac{\partial^2 f}{\partial y^2}\end{aligned}$$

3.3 More interesting examples

Example 7: Axially symmetric functions $V(x,y)$ satisfying Laplace's equation $V_{xx}+V_{yy}=0$.

If we rewrote the function $V(x,y)$ in to polar coordinates, from the previous example we know the form of $\frac{\partial^2 V}{\partial r^2}$. We can similarly compute $\frac{\partial^2 V}{\partial \theta^2}$. The details are on the **handout,** we find that

$$\partial_x^2 V + \partial_y^2 V = \partial_r^2 V + \frac{1}{r}\partial_r V + \frac{1}{r^2}\partial_\theta^2 V = 0.$$

If axially symmetric then $V = V(r)$ only and it reduces to an **o.d.e.**

$$V_{rr} + \frac{1}{r}V_r = 0.$$

Let $p = V_r$ then $rp_r + p = 0$ so that $rp = c$ and $V_r = \frac{c}{r}$ so that

$$V = c\log(\sqrt{x^2+y^2}) + d.$$

The important point here is that sometimes it is possible to use coordinates where the problem is reduced to an o.d.e.

Example 8: Find the general solution $f(x,t)$ of $3f_x - f_t = 0$.

This is a typical sort of equation that arises when there is a traveling wave as we'll see. Change variables to $s(x,t) = x+3t$ and $r(x,t) = x-3t$, then the change of variables rule gives

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial f}{\partial s}\frac{\partial s}{\partial x} + \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} = f_s + f_r \\ \frac{\partial f}{\partial t} &= \frac{\partial f}{\partial s}\frac{\partial s}{\partial t} + \frac{\partial f}{\partial r}\frac{\partial r}{\partial t} = 3f_s - 3f_r. \end{aligned}$$

Then $3f_x - f_t = 6f_r = 0$. So we see that the solution is $f = g(s) = g(x+3t)$, for arbitrary g. If we are given an initial condition such as $f(x,0) = \sin(x)$ then $f(x,t) = \sin(x+3t)$. This is a wave moving left at speed 3.

Example 9: The wave equation $u_{xx} - c^{-2}u_{tt} = 0$.

This is the type of equation describing the string, and many other systems. The way to solve this problem can be seen by rewriting it (using $a^2 - b^2 =$

$(a+b)(a-b)$ – note we can use this relation for differential operators too!) as

$$(\partial_x - c^{-1}\partial_t)(\partial_x + c^{-1}\partial_t)\, u = 0.$$

Suppose $s = x + ct$ and $r = x - ct$. Then again following example 8 by the chain rule

$$\begin{aligned}
\partial_x u &= \partial_s u \partial_x s + \partial_r u \partial_x r = \partial_s u + \partial_r u \\
\partial_t u &= \partial_s u \partial_t s + \partial_r u \partial_t r = c\partial_s u - c\partial_r u
\end{aligned}$$

Then $(\partial_x - c^{-1}\partial_t) = 2\partial_r$ and $(\partial_x + c^{-1}\partial_t) = 2\partial_s$ so the equation becomes

$$u_{rs} = 0.$$

Integrating as in example 3 gives $u = f(r) + g(s) = f(x+ct) + g(x-ct)$. This is d'Alembert's solution.

This is the sum of a left moving wave with speed c and a right moving wave. So if this is the general solution the oscillating string must be a solution like this.

To see how a standing oscillating mode can be realized like this consider the

$$u = u_0\, [\sin(\frac{n\pi}{L}(x+ct)) + \sin(\frac{n\pi}{L}(x-ct))].$$

Using $\sin(A+B) + \sin(A-B) = 2\cos B\, \sin A$ we get

$$u = 2u_0 \cos(\frac{n\pi c}{L} t)\, \sin(\frac{n\pi}{L} x).$$

4 Taylor expansions (Riley 5.7)

For a function of a single variable Taylor's theorem allows us to expand about a point x_0. Writing $x = x_0 + \Delta x$

$$F(x) = F(x_0 + \Delta x) = F(x_0) + \Delta x\, F'(x_0) + \frac{\Delta x^2}{2!} F''(x_0) + \ldots$$

Can also write it as an *operator equation* using the fact that $e^A = 1 + A + \frac{A^2}{2!} + \ldots$

$$F(x_0 + \Delta x) = e^{\Delta x \frac{d}{dx}} F(x)|_{x=x_0}.$$

This last equation makes it obvious how to generalize; a function of two variables expanded about $(x, y) = (x_0, y_0)$ can be found by first expanding about $x = x_0$ and then about $y = y_0$;

$$\begin{aligned} f(x,y) = f(x_0 + \Delta x, y_0 + \Delta y) &= e^{\Delta x \partial_x} e^{\Delta y \partial_y} f(x,y)|_{x=x_0, y=y_0} \\ &= f + \Delta x\, f_x| + \Delta y\, f_y| + \frac{1}{2}(\Delta x^2 f_{xx} + 2\Delta x \Delta y f_{xy} + \Delta y^2 f_{yy})| + ... \end{aligned}$$

where the vertical line | indicates setting $x \to x_0$ and $y \to y_0$.

Example 1: $\cos(x + y)$ **about (0,0) up to and including quadratics**

In above have $x_0 = 0 = y_0$ and $\Delta x = x$ and $\Delta y = y$: have

$$\begin{aligned} f &= \cos(x + y) \\ f_x = f_y &= -\sin(x + y) \\ f_{xx} = f_{xy} = f_{yy} &= -\cos(x + y) \end{aligned}$$

Then

$$\begin{aligned} f(x,y) &= f + x f_x + y f_y + \frac{1}{2}(x^2 f_{xx} + 2xy f_{xy} + y^2 f_{yy})|_{0,0} \\ &= 1 - \frac{1}{2}(x + y)^2 + ... \end{aligned}$$

Note this is slightly trivial since could have just expanded $z = x + y$

Example 2: $f(x, y) = \log(x + 2y)$,**find the Taylor expansion about (1,0)**

In earlier notation, $x_0 = 1$, $y_0 = 0$, $\Delta x = x - 1$ and $\Delta y = y$. Have

$$\begin{aligned}
f| &= \log(x+2y)| = 0 \\
f_x| &= \frac{1}{x+2y}| \ = 1 \\
f_{xx}| &= -\frac{1}{(x+2y)^2}| = -1 \\
f_y| &= \frac{2}{x+2y}| \ = 2 \\
f_{yy}| &= -\frac{4}{(x+2y)^2}| \ = -4 \\
f_{xy}| &= -\frac{2}{(x+2y)^2}| \ = -2
\end{aligned}$$

So that

$$\begin{aligned}
f(x, y) &= (x-1) + 2y - \frac{1}{2}((x-1)^2 + 4(x-1)y + 4y^2) + ... \\
&= z - \frac{z^2}{2} + ... \quad [z = x + 2y - 1]
\end{aligned}$$

Again could have done this more simply.

5 Critical points

Want to find where local maxima minima or saddle points are. A critical point is a point at which both $f_x = 0$ and $f_y = 0$.

Example 1: $f(x,y) = x^2 + y^2$**.**

The graph of $z = f(x,y)$ is a parabolic cylinder.

$$f_x = 2x \ f_y = 2y$$

The point (0,0) is the only critical point.

Example 2: $f(x,y) = x^2 - y^2$**.**

The graph of $z = f(x,y)$ is a saddle.

$$f_x = 2x \ f_y = -2y$$

The point (0,0) is the only critical point. Increases away from origin for fixed y but decreases for fixed y.

5.1 Local maxima and minima

Defn: A point (x_0, y_0) is said to be a local *maximum* if $f(x_0, y_0) > f(x, y)$ for all points (x, y) in a sufficiently small neighbourhood surrounding (x_0, y_0).

Defn: A point (x_0, y_0) is said to be a local *minimum* if $f(x_0, y_0) < f(x, y)$ for all points (x, y) in a sufficiently small neighbourhood surrounding (x_0, y_0).

We can use the Taylor expansion about (x_0, y_0) to tell us about the nature of the point there. To simplify things call $\Delta x = x - x_0$ and $\Delta y = y - y_0$ and to simplify notation we write

$$\begin{aligned} f_{xx} &= f_{xx}(x_0, y_0) \\ f_{xy} &= f_{xy}(x_0, y_0) \\ f_{yy} &= f_{yy}(x_0, y_0). \end{aligned}$$

Then we can write

$$f(x,y) = f(x_0,y_0)+\Delta x\, f_x(x_0,y_0)+\Delta y\, f_y(x_0,y_0)+\frac{1}{2}(\Delta x^2 f_{xx}+2\Delta x\Delta y f_{xy}+\Delta y^2 f_{yy})+\ldots$$

A necessary condition for a maximum, minimum or saddle is that $f_x = f_y = 0$.

The test for what sort of critical point it is is let $M = f_{xx}f_{yy} - (f_{xy})^2$

- If $M > 0$ and $f_{xx} > 0$ then local minimum
- If $M > 0$ and $f_{xx} < 0$ then local maximum
- If $M < 0$ then saddle point
- If $M = 0$ then inconclusive

Proof: the function near x_0, y_0 is

$$\begin{aligned} f(x_0+\Delta x, y_0+\Delta y) &= f(x_0,y_0) + \frac{1}{2f_{xx}}(\Delta x^2 f_{xx}^2 + 2\Delta x\Delta y\, f_{xy}f_{xx} + \Delta y^2 f_{yy}f_{xx}) + \ldots \\ &= f(x_0,y_0) + \frac{1}{2f_{xx}}((\Delta x\, f_{xx} + \Delta y\, f_{xy})^2 + \Delta y^2 M) + \ldots . \end{aligned}$$

If $f_{xx} > 0$ and $M > 0$ then $f(x,y)-f(x_0,y_0) > 0$ and $f(x_0,y_0)$ is a minimum. If $f_{xx} < 0$ then the reverse is true. If $M < 0$ then for some values of $\Delta x, \Delta y$, $f(x,y) - f(x_0,y_0)$ is positive and for others it's negative.

5.2 Examples

Example 1: $f(x,y) = x^2 + y^2$

The graph of $z = f(x,y)$ is a parabolic cylinder.

$$f_{xx} = 2,\ f_{yy} = 2,\ f_{xy} = 0$$

And then $M = 4$ and $f_{xx} = 2$, it works!

Example 2: $f(x,y) = x^2 - y^2$

Should find that it is a saddle

$$f_{xx} = 2,\ f_{yy} = -2,\ f_{xy} = 0$$

And then $M = -4$, it works!

Example 3: Find and classify the critical points of $f(x, y) = -x^3 + 4xy - 2y^2 + 1$

First find where $f_x = f_y = 0$.

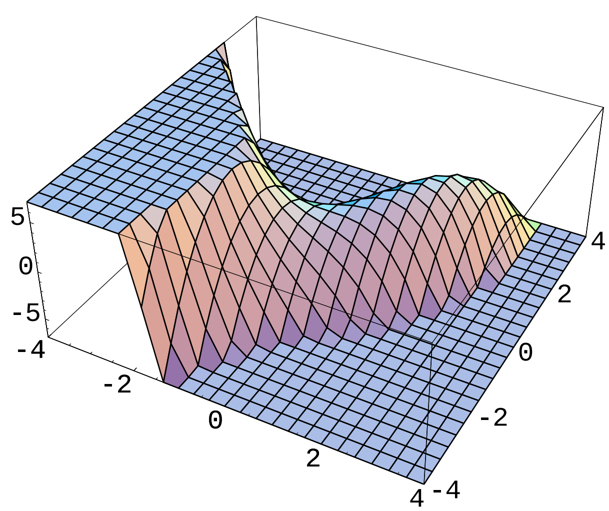
$$\begin{aligned} f_x &= -3x^2 + 4y \\ f_y &= 4x - 4y \\ f_{xx} &= -6x \\ f_{yy} &= -4 \\ f_{xy} &= 4 \end{aligned}$$

Solving these gives $x = y$ and $x(-3x + 4) = 0$ so have

$$(x, y) = (0, 0) \textit{ or } (\frac{4}{3}, \frac{4}{3})$$

which I'll label A and B.

- At A we have $f_{xx} = 0$, $f_{yy} = -4$, $f_{xy} = 4$ so $M = -4$ and A is a saddle
- At B we have $f_{xx} = -8$,$f_{yy} = -4$, $f_{xy} = 4$ so $M = 16 > 0$. Also $f_{xx} < 0$ so that B is a maximum.



Example 4: Find and classify the critical points of $f(x,y) = x^2 - y^2 + y^4 + x^2y^2$

First find where $f_x = f_y = 0$.

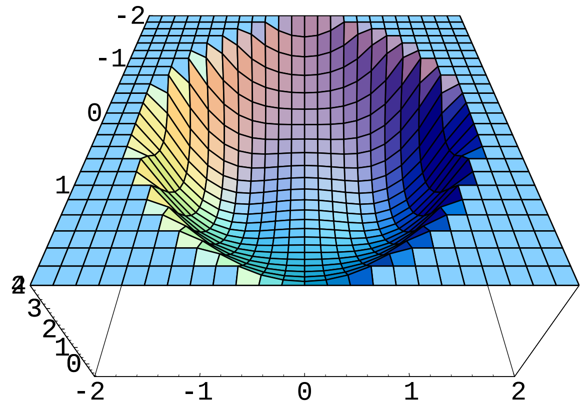
$$\begin{aligned} f_x &= 2x(1+y^2) \\ f_y &= 2y(x^2+2y^2-1) \\ f_{xx} &= 2(1+y^2) \\ f_{yy} &= 2(x^2+6y^2-1) \\ f_{xy} &= 4xy \end{aligned}$$

Solving $f_x = f_y = 0$ gives $x = 0$ and $y = 0, \pm\frac{1}{\sqrt{2}}$ (since $y^2+1 \geq 1$ then f_x can only give $x = 0$) so have

$$(x,y) = (0,0) \ or \ (0, \frac{1}{\sqrt{2}}) \ or \ (0, -\frac{1}{\sqrt{2}})$$

which I'll label A, B and C.

- At A we have $f_{xx} = 2$, $f_{yy} = -2$, $f_{xy} = 0$ so $M = -4$ and A is a saddle
- At B and C we have $f_{xx} = 3$,$f_{yy} = 4$, $f_{xy} = 0$ so $M = 12 > 0$. Also $f_{xx} > 0$ so that B, C are minima.



Example 5: A rectangular box *open at the top* is to have a volume of $32m^3$. What are the dimensions in order to make the surface area as small as possible?

First write the expressions for the volume and surface area if the base width height are x, y, z;

$$\begin{aligned} V &= xyz \\ S &= xy + 2xz + 2yz. \end{aligned}$$

For a given x, y the volume determines $z = \frac{V}{xy}$ where $V = 32m^3$. Then

$$S = xy + 2\frac{V}{y} + 2\frac{V}{x}.$$

Extrema of S will be found where $S_x = S_y = 0$;

$$\begin{aligned} S_x &= 0 = & y - 2\frac{V}{x^2} \\ S_y &= 0 = & x - 2\frac{V}{y^2} \end{aligned}$$

Solving this gives $x = y$ and then $x^3 = 2V$ or $x = y = (2V)^{\frac{1}{3}} = 4m$, $z = \frac{1}{2}(2V)^{\frac{1}{3}} = 2m$.

This should be a minimum, so need $S_{xx} = \frac{4V}{x^3} = 2$, $S_{yy} = \frac{4V}{y^3} = 2$, $S_{xy} = 1$, so that $M = S_{xx}S_{yy} - (S_{xy})^2 > 0$ and $S_{xx} > 0$ - i.e. it is indeed a local minimum.

6 Double integrals (Riley 6.1)

Consider solving a p.d.e. directly by integration such as

$$u_{xy} = 2x^2y^3 + 2y$$

We integrate with respect to y with x constant, and then w.r.t. x keeping y constant. The first step gives

$$u_x = \frac{x^2y^4}{2} + y^2 + F(x)$$

and the second gives

$$u = \frac{x^3y^4}{6} + xy^2 + f(x) + g(y).$$

where $f_x(x) = F(x)$.

These were both *indefinite integrals.* But we could consider doing definite integrals instead. Consider putting limits in the first integration of $0, 2$. Then

$$\begin{aligned} \int_0^2 (2x^2y^3 + 2y)dy &= [\frac{x^2y^4}{2} + y^2]_0^2 \\ &= 8x^2 + 4. \end{aligned}$$

Now consider for example doing the second integral with limits $0, 1$:

$$\begin{aligned} \int_0^1 (8x^2 + 4)dx &= [8\frac{x^3}{3} + 4x]_0^1 \\ &= \frac{8}{3} + 4 = \frac{20}{3}. \end{aligned}$$

This is an example of a repeated integral. It can be written as a single expression

$$\int_0^1 \int_0^2 (2x^2y^3 + 2y)\, dy\, dx = \frac{20}{3}$$

with the convention that the ordering of $dydx$ tells us to do the y integral first. The integral has an interpretation as the "volume under the surface given by $f(x, y)$" which I'll discuss in a moment. For the moment simply note that the area over which we integrated is the rectangle $\{x, y|\, x \in [0, 1],\ y \in [0, 2]\}$.

The limits on the first y-integration could have depended on x. For example we could have taken limits $0 \leq y \leq x$. Then we would have

$$\int_0^1 \int_0^x (2x^2y^3 + 2y)\, dy\, dx.$$

The y integration gives

$$\begin{aligned} \int_0^x (2x^2y^3 + 2y) dy &= [\frac{x^2y^4}{2} + y^2]_0^x \\ &= \frac{x^6}{2} + x^2 \end{aligned}$$

and then the second integral with limits $0, 1$ gives

$$\begin{aligned} \int_0^1 (\frac{x^6}{2} + x^2) dx &= [\frac{x^7}{14} + \frac{x^3}{3}]_0^1 \\ &= \frac{1}{14} + \frac{1}{3} = \frac{17}{42}. \end{aligned}$$

The region over which we integrated is now a triangle with hypotenuse given by the line $y = x$. I'll come to more complicated regions of integration in a moment.

6.1 Interpretation as volume under a surface

Recap of the one-dimensional (Riemann) integration

Suppose we want to calculate the area under a curve $f(x)$ between $x = a$ and $x = b$: that area is

$$S = \int_a^b f(x) dx.$$

To approximate this area we can divide the interval into N lengths $\Delta x_1, \Delta x_2 \ldots \Delta x_N$. Let x_i be the point midway along the i'th interval. Then the area is approximately given by the sum of the area of all the rectangles of height $f(x_i)$ and width Δx_i:

$$S \approx \sum_{i=1}^{N} f(x_i)\, \Delta x_i.$$

To find the integral we take the limit as $N \to \infty$

$$S = \lim_{N \to \infty} (\sum_{i=1}^{N} f(x_i)\, \Delta x_i) := \int_a^b f(x) dx.$$

Extension to two-dimensional (Riemann) integration

Suppose we want to calculate the volume under a surface whose height is given by $f(x, y)$. Let the region in the x, y plane that we want to integrate over be R so that we might guess that

$$V = \iint_R f(x, y)\, dA.$$

To see that this is indeed the volume, approximate the volume by dividing the integration region into N areas $\Delta A_1, \Delta A_2 \ldots \Delta A_N$. Let x_i, y_i be the point "in the middle" of the i'th area. Then the volume is approximately given by the sum of the volumes of all the columns of height $f(x_i, y_i)$ and base area ΔA_i:

$$V \approx \sum_{i=1}^{N} \Delta A_i\, f(x_i, y_i).$$

The double integral is defined as the limit as $N \to \infty$ so that

$$V \;= \textstyle\iint_R f(x, y)\, dA \;=\; \displaystyle\lim_{N\to\infty} (\sum_{i=1}^{N} \Delta A_i f(x_i, y_i).)$$

6.2 Explicit example: integration over rectangles

So far, so formal. How do we see that the limit above is the same thing as the earlier double integrations? Consider the earlier example where R was the rectangle given by $0 \leq x \leq 1$ and $0 \leq y \leq 2$. The integral we wish to evaluate should be the volume under the surface of height $f(x, y) = 2x^2y^3 + 2y$. The area summation in this case is most easily done by summing the area first in the y direction and then the x direction. That is we introduce a *double sum* over i,j and then define the area elements as small rectangles $\Delta A_{ij} = \Delta x_i \Delta y_j$;

$$V \approx \sum_{i=1}^{n} \sum_{j=1}^{m} \Delta x_i \Delta y_j\, f(x_i, y_j)$$

with $N = nm$. But another way to think of this is that for each x_i we are evaluating the area under the curve $f(x_i, y)$ at x_i given by the single integration limit

$$S(x_i) \;=\; \lim_{m\to\infty} \sum_{j=1}^{m} f(x_i, y_j)\Delta y_j = \int_0^2 f(x_i, y) dy$$

and then summing the volumes of all the slices of thickness Δx_i given by

$$\begin{aligned} V &= \lim_{n\to\infty} \sum_{i=1}^{n} \Delta x_i S(x_i) \\ &= \int_0^1 S(x) dx \\ &= \int_0^1 \int_0^2 f(x,y)\, dy dx \end{aligned}$$

where we simply twice utilized our definition of single integration as a limit.

- Note that in this particular case we can **do the sum in either order**, but with integration over more complicated areas (i.e. not rectangles) you have to *pay attention to the limits.*

- **An important example** of double integration region is when $f(x,y) = 1$; the "volume" under this surface is clearly just **the area of the integration region**;
$$V = \iint_R dA = A.$$

Example 1: Find $\iint_R \frac{1}{(x+y+1)^2} dA$ where R is the square defined by $0 \le x \le 1$ and $0 \le y \le 1$.

We can do the integral

$$\begin{aligned} \int_0^1 \int_0^1 \frac{1}{(x+y+1)^2} dx dy &= \int_0^1 \left[-\frac{1}{(x+y+1)} \right]_0^1 dy \\ &= \int_0^1 \left(\frac{1}{(y+1)} - \frac{1}{(y+2)} \right) dy \\ &= \left[\log \left(\frac{y+1}{y+2} \right) \right]_0^1 \\ &= \log \frac{2.2}{3.1} = \log \frac{4}{3} \end{aligned}$$

Example 2: Find $\iint_R x\sin(xy)dA$ where R is the rectangle defined by $\frac{\pi}{2} \leq x \leq \pi$ and $0 \leq y \leq 1$.

In this case it is easier to do the y integral first (as it's just like integrating a $\sin y$ function). We can do the integral

$$\begin{aligned}
\int_{\frac{\pi}{2}}^{\pi}\int_0^1 x\sin(xy)dydx &= \int_{\frac{\pi}{2}}^{\pi}\left[-\cos(xy)\right]_{y=0}^{y=1}dx \\
&= \int_{\frac{\pi}{2}}^{\pi}(1-\cos x)\,dx \\
&= \left[x-\sin x\right]_{\frac{\pi}{2}}^{\pi} \\
&= \frac{\pi}{2}+1.
\end{aligned}$$

6.3 Double integrals over other regions

Suppose that instead the region R is defined by

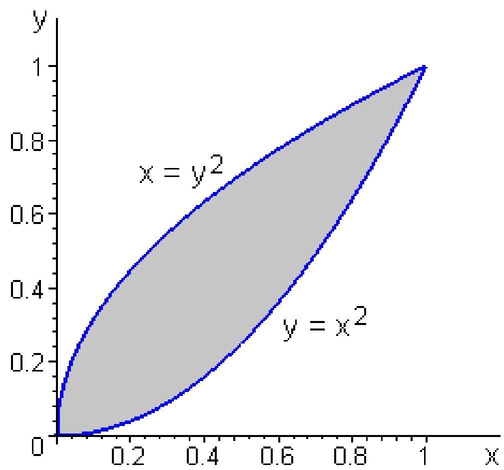
$$\begin{aligned}
a &\leq x \leq b \\
u(x) &\leq y \leq v(x)
\end{aligned}$$

where u, v are some functions. Most cases of interest *could* be defined this way although some are written more naturally in different coordinate systems (e.g. polar) than Cartesian x, y. Integration becomes

$$\iint_R f(x,y)dA = \int_a^b\int_{u(x)}^{v(x)} f(x,y)dy\,dx.$$

Note that the interpretation as the limit of sums is unchanged from earlier, but we cannot not now easily change the order of integration as the limits are explicitly x dependent. In the figure below I show the region of integration for the integral

$$\int_0^1\int_{x^2}^{\sqrt{x}} f(x,y)\,dydx$$



Example 1a: Evaluate $\iint_R xy\,dA$ where R is the finite region enclosed by the line $y = x$ and the curve $y = x^2$.

The two curves intersect at $x = 0$ and $x = 1$. The region of integration is therefore defined by

$$\begin{aligned} 0 &\leq x \leq 1 \\ x^2 &\leq y \leq x \end{aligned}$$

(Note that $x \geq x^2$ in this region). The way the problem is posed we can obviously do the y integral first:

$$\begin{aligned} \int_0^1 \int_{x^2}^x xy\,dydx &= \int_0^1 \left[\frac{1}{2}xy^2\right]_{y=x^2}^{y=x} dx \\ &= \int_0^1 \left(\frac{x^3}{2} - \frac{x^5}{2}\right) dx \\ &= \frac{1}{2}\left[\frac{x^4}{4} - \frac{x^6}{6}\right]_0^1 \\ &= \frac{1}{24} \end{aligned}$$

Example 1b: A first look at changing the order of integration:

Suppose we wish to do the x integration first. We should get the same answer. The region of integration must first be rewritten as

$$\begin{aligned} 0 \ &\leq \ y \leq 1 \\ y \ &\leq \ x \leq \sqrt{y} \end{aligned}$$

(Note that $\sqrt{y} \geq y$ in this region). We should now do the x integral first:

$$\begin{aligned} \int_0^1 \int_y^{\sqrt{y}} xy \, dx dy &= \int_0^1 \left[\frac{1}{2} y x^2 \right]_{x=y}^{x=\sqrt{y}} dy \\ &= \int_0^1 \left(\frac{y^2}{2} - \frac{y^3}{2} \right) dy \\ &= \frac{1}{2} \left[\frac{y^3}{3} - \frac{y^4}{4} \right]_0^1 \\ &= \frac{1}{24} \quad (\textit{phew}) \end{aligned}$$

Example 2: Find the area of the finite region R in the plane bounded by the the finite region enclosed by the line $x = y^2$ and the lines $x = 3$ and $y = \pm 1$.

First note that to find the area we can just do the double integral with $f(x, y) = 1$. Inspecting the region of integration, the easiest is to do the x integration first (otherwise we would have to split the x-integration into two regions). The region of integration is therefore defined by

$$\begin{aligned} y^2 \ &\leq \ x \leq 3 \\ -1 \ &\leq \ y \leq 1 \end{aligned}$$

We have to now do the x integral first:

$$\begin{aligned}\int_{-1}^{1}\int_{y^2}^{3} dxdy &= \int_{-1}^{1} [x]_{x=y^2}^{x=3}\, dy \\ &= \int_{-1}^{1} \left(3-y^2\right) dy \\ &= 2\left[3y-\frac{1}{3}y^3\right]_0^1 \\ &= 2.\frac{8}{3} = \frac{16}{3}\end{aligned}$$

On the 3rd line I used the symmetry of the region around the $x-$axis, changed the integration region from [-1,1] to [0,1] and multiplied by 2.

Example 3: Find the area of the bounded region R determined by the curves in the plane $y = x^2/4$ and $2y - x - 4 = 0$.

The bounded region is everywhere below where the straight line intersects the parabola. Again inspecting the region of integration, the easiest is to do the y integration first. The intersection points are when

$$\frac{x^2}{2} = x + 4$$

which has solutions at $x = -2, 4$. The region of integration is therefore defined by

$$\begin{aligned}\frac{x^2}{4} &\le y \le \frac{1}{2}(x+4) \\ -2 &\le x \le 4\end{aligned}$$

We have to now do the y integral first:

$$\begin{aligned}\int_{-2}^{4}\int_{x^2/4}^{(x+4)/2} dydx &= \int_{-2}^{4} [y]_{y=x^2/4}^{y=(x+4)/2}\, dx \\ &= \int_{-2}^{4} \frac{1}{4}\left(2x+8-x^2\right) dx \\ &= \frac{1}{4}\left[x^2+8x-\frac{1}{3}x^3\right]_{-2}^{4} \\ &= \frac{1}{4}(16+32-\frac{64}{3}-4+16-\frac{8}{3}) = \frac{36}{4} = 9.\end{aligned}$$

6.4 More on changing the order of integration

Sometimes it may be preferable to carry out the integration in a different order, but as we have seen you need to be a little careful with the limits.

Example 1: Find the integral $I = \int_0^1 \int_y^1 \cos(\frac{\pi}{2}x^2)\, dxdy$**.**

In this case the x integral is not easy (it gives something called the Fresnel integral). The easiest route is to do the y integral first. The integration region is everywhere above the line $y = x$ up to $x = 1$. The region of integration defined above is

$$\begin{aligned} y &\leq x \leq 1 \\ 0 &\leq y \leq 1 \end{aligned}$$

which we can rewrite as

$$\begin{aligned} 0 &\leq y \leq x \\ 0 &\leq x \leq 1 \end{aligned}$$

We can now do the y integral first:

$$\begin{aligned} \int_0^1 \int_0^x \cos(\frac{\pi}{2}x^2) dydx &= \int_0^1 \left[y \cos(\frac{\pi}{2}x^2)\right]_{y=0}^{y=x} dx \\ &= \int_0^1 \left(x \cos(\frac{\pi}{2}x^2)\right) dx \\ &= \left[\frac{1}{\pi}\sin(\frac{\pi}{2}x^2)\right]_0^1 \\ &= \frac{1}{\pi}. \end{aligned}$$

Example 2: Find the integral $I = \int_0^1 \int_{\sqrt{y}}^1 \exp(x^3)\, dxdy$**.**

In this case the x integral is not easy (it gives something depending on what are called incomplete Gamma functions). The easiest route is to do the y integral first. The integration region is everywhere above the line $y = x$ up to $x = 1$. The region of integration defined above is

$$\begin{aligned} \sqrt{y} &\leq x \leq 1 \\ 0 &\leq y \leq 1 \end{aligned}$$

which we can rewrite as

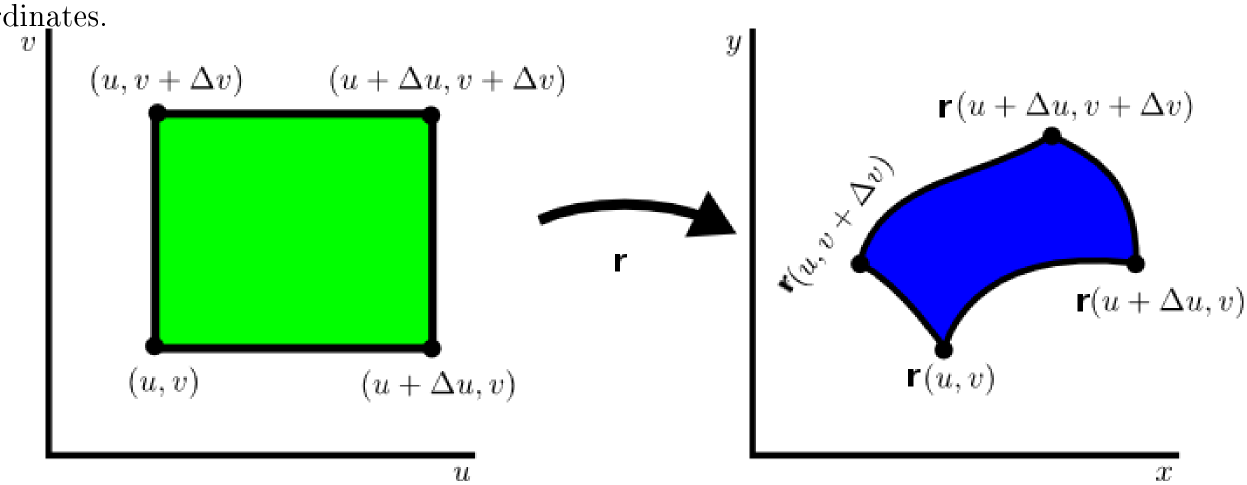
$$\begin{aligned} 0 &\leq y \leq x^2 \\ 0 &\leq x \leq 1 \end{aligned}$$

We can now do the y integral first:

$$\begin{aligned} \int_0^1 \int_0^{x^2} \exp(x^3) dy dx &= \int_0^1 \left[y \exp(x^3)\right]_{y=0}^{y=x^2} dx \\ &= \int_0^1 \left(x^2 \exp(x^3)\right) dx \\ &= \left[\frac{1}{3} \exp(x^3)\right]_0^1 \\ &= \frac{(e-1)}{3}. \end{aligned}$$

6.5 Change of variables in double integration (Riley 6.4.1)

We now know how to write $\iint_R f(x,y)dA$ as a double integral in Cartesian co-ordinates $\iint_R f(x,y)dA = \int_a^b \int_{u(x)}^{v(x)} f(x,y)dydx$. The important point here is that the measure $dA = dx\,dy$. What happens if we use alternative co-ordinates (eg Polar coordinates etc.)? To answer this we ask what happens to the area element (an infinitesimally small piece of area) in different co-ordinates.



A small area in (u, v) space is mapped into physical space $\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j}$. As Δu and Δv become small, this blue area becomes closer and closer to a parallelogram. The edges of this small parallelogram are given by the vectors $\mathbf{a}, \mathbf{b}$:

$$\begin{aligned}
\mathbf{a} = &\mathbf{r}(u + \Delta u, v) - \mathbf{r}(u, v) &\approx& \frac{\partial \mathbf{r}}{\partial u}\Delta u \\
\mathbf{b} = &\mathbf{r}(u, v + \Delta v) - \mathbf{r}(u, v) &\approx& \frac{\partial \mathbf{r}}{\partial v}\Delta v
\end{aligned}$$

The area of the small parallelogram with edges $\mathbf{a}$ and $\mathbf{b}$ is given by $\Delta A = |\mathbf{a} \times \mathbf{b}| = |\left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right)| \Delta u \Delta v = |x_u y_v - x_v y_u| \Delta u \Delta v$. Taking the limit to zero we get

$$dA = |x_u y_v - x_v y_u|\, du\, dv = J\, du\, dv.$$

Here $J = |x_u y_v - x_v y_u|$is known as the Jacobian and is sometimes denoted

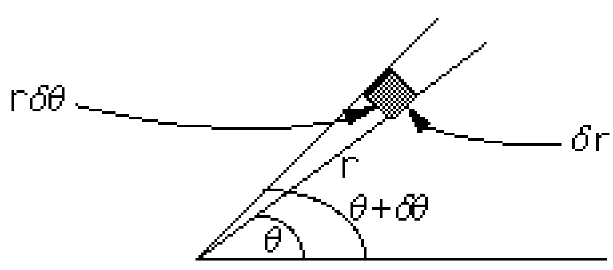
$$\frac{\partial(x, y)}{\partial(u, v)} = J = |x_u y_v - x_v y_u|.$$

So the integral $\iint_R f(x, y) dA = \iint_R f(x, y) dx dy$ can be found in the new co-ordinates as

$$\iint_R f(x, y) dA = \iint_R f(x, y) dx dy = \iint_R f(x(u, v), y(u, v)) \frac{\partial(x, y)}{\partial(u, v)} du dv = \iint_R f(x(u, v), y(u, v)) |x_u$$

6.6 Use of polar coordinates

For some problems with axial symmetry it is easier to use polar coordinates. We cover the x, y plane with a polar grid and our area elements ΔA_i become wedges at r, θ as shown;



The area of an element of width Δr and $r\Delta\theta$ is given by $\Delta A = r\Delta r\,\Delta\theta$. In addition of course $f(x,y) \to f(r\cos\theta, r\sin\theta)$ so that we have the theorem

$$\iint_R f(x,y)dA = \iint_R f(r\cos\theta, r\sin\theta)\, rdr\, d\theta$$

Compare this with the general change of variables given above. Here $x = r\cos\theta$, $y = r\sin\theta$

$$J = \frac{\partial(x,y)}{\partial(r,\theta)} = |x_r y_\theta - x_\theta y_r| = |\cos\theta(r\cos\theta) - (-r\sin\theta)\sin\theta| = r$$

and so again from the previous subsection the area element $dA = J\,dr\,d\theta = r\,dr\,d\theta$.

Example 1: Evaluate $I = \iint_R \sin(x^2+y^2)\,dA$ where R is the circular disc $x^2+y^2 \le \pi$.

Clearly in this case there is a circular symmetry we should be taking advantage of. The region of integration defined above is

$$\begin{aligned} 0 \;&\le\; r \;\le\; \sqrt{\pi} \\ 0 \;&\le\; \theta \;\le\; 2\pi \end{aligned}$$

Using the theorem above we can write

$$\begin{aligned} I &= \int_0^{\sqrt{\pi}} \int_0^{2\pi} \sin(r^2)\, r\, d\theta\, dr \\ &= 2\pi \int_0^{\sqrt{\pi}} \sin(r^2)\, r\, dr \\ &= \left[-\pi\cos(r^2)\right]_0^{\sqrt{\pi}} \\ &= \pi(1-\cos(\pi)) = 2\pi. \end{aligned}$$

Example 2: Evaluate $I = \iint_R 3xy^2\,dA$ where R is the semi-circular disc $x^2+y^2 \le 1$ and $x \ge 0$.

First note that the line $x = 0$ in polar coordinates runs along $\theta = \pm\frac{\pi}{2}$. The region of integration defined above is

$$\begin{aligned} 0 \;&\le\; r \;\le\; 1 \\ -\frac{\pi}{2} \;&\le\; \theta \;\le\; \frac{\pi}{2} \end{aligned}$$

Using the theorem above we can write

$$\begin{aligned}
I &= \int_0^1 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 3r^3 \cos\theta \sin^2\theta \, r d\theta \, dr \\
&= 2 \int_0^1 \int_0^1 3r^4 \sin^2(\theta) \, d(\sin\theta) \, dr \\
&= 2 \int_0^1 r^4 \, dr \\
&= \left[\frac{2}{5} r^5 \right]_0^1 \\
&= \frac{2}{5}.
\end{aligned}$$

Example 3: Evaluate $I = \int_{-\infty}^{\infty} e^{-x^2} \, dx$

This is a famous integral (the Gaussian integral) whose answer is known to be $\sqrt{\pi}$. To do it first construct the square;

$$\begin{aligned}
I^2 &= \int_{-\infty}^{\infty} e^{-x^2} \, dx \int_{-\infty}^{\infty} e^{-y^2} \, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} \, dx \, dy
\end{aligned}$$

This is the integration over the entire plane. Using polar coordinates it becomes simpler;

$$\begin{aligned}
I^2 &= \int_0^{\infty} \int_0^{2\pi} e^{-r^2} d\theta \, r dr \\
&= 2\pi \int_0^{\infty} e^{-r^2} \, r dr \\
&= \left[-\pi e^{-r^2} \right]_0^{\infty} = \pi
\end{aligned}$$

so $I = \sqrt{\pi}$.

7 Extension to triple integration

Many problems require the evaluation of a "bulk" property such as mass of a 3-dimensional object, by integrating over the entire volume. Suppose we want to find the mass of a star for example, of density $\rho(x, y, z)$. The mass of a small volume element with sides Δx, Δy, Δz will be

$$\Delta M = \rho(x, y, z)\Delta x \Delta y \Delta z$$

Extending the two dimensional case then, we can subdivide the volume into N^3 small elements, and sum. The total mass can be written as the triple integral

$$M = \iiint_V \rho \, dxdydz$$

or if we wish to avoid specifying coordinates

$$M = \iiint_V \rho \, dV$$

where V is the integration volume.

Example 1: A wedge occupies the region V given by $0 \le x \le 2$, $0 \le y \le 1$ and $0 \le z \le 1 - y$. The wedge is made of material of density $\rho = 12\, xy\, g/cm^3$. What is its mass?

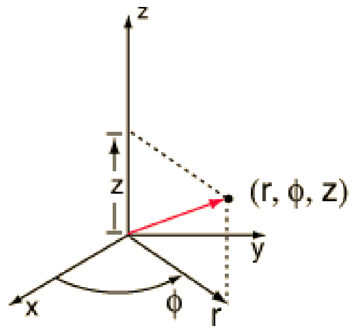
We have

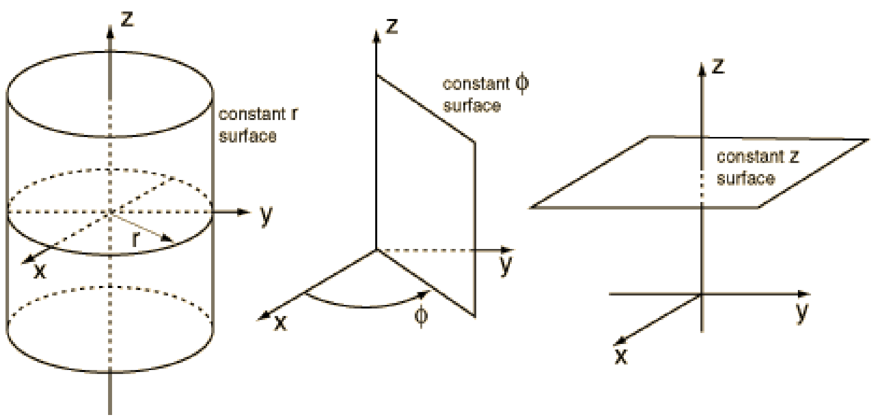
$$
\begin{aligned}
M &= \iiint_V \rho \, dxdydz \\
&= 12 \int_0^2 \int_0^1 \int_0^{1-y} xy \; dzdydx \\
&= 12 \int_0^2 \int_0^1 xy \, [z]_0^{1-y} \; dydx \\
&= 12 \int_0^2 \int_0^1 xy(1-y) \; dydx \\
&= 12 \int_0^2 x \left[\frac{y^2}{2} - \frac{y^3}{3} \right]_0^1 dx \\
&= 2 \int_0^2 xdx = [x^2]_0^2 = 4\, gm
\end{aligned}
$$

Note that implicit in this example were the units of length in the limits (e.g. 2cm, $(1-y)$cm etc).

7.1 Cylindrical polar coordinates

As for double integration, there are often cases were different coordinate systems are preferable. Cylindrical polar coordinates are useful in cases with axial symmetry, usually involving pipes, cylinders, annuluses etc. They are an extension to polars where we essentially just "add a $z-$coordinate". It is illustrated in the figure, (to tie in with convention I'll now call the angular coordinate ϕ)





If we wish to integrate over the volume the volume elements have sides Δr, $r\Delta\phi$ and Δz so that

$$\iiint_V f(x,y,z)\,dxdydz \quad = \quad \iiint_V f(r\cos\phi, r\sin\phi, z)\,rdrd\phi dz$$

Example 2: Find the volume V of the finite region bounded by the cylinder $x^2+y^2=1$, and the planes $z=0$ and $z=x+3$.

We have to find $V=\iiint_V dxdydz$ for this volume. Given the cylindrical symmetry of at least part of the problem we might try doing $V=\iiint_V rdrd\phi dz$ instead. Clearly the tricky bit is the z limit. This will become

$$0 \le z \le r\cos\phi + 3$$

and should be evaluated first. That is we do

$$\begin{aligned}
V &= \int_0^1 \int_0^{2\pi} \int_0^{r\cos\phi+3} r dz d\phi dr \\
&= \int_0^1 \int_0^{2\pi} r\,[z]_0^{r\cos\phi+3}\, d\phi dr \\
&= \int_0^1 \int_0^{2\pi} r(r\cos\phi+3)\, d\phi dr \\
&= \int_0^1 r\,[r\sin\phi+3\phi]_0^{2\pi}\, dr \\
&= \int_0^1 6\pi r dr \\
&= 3\pi[r^2]_0^1 = 3\pi.
\end{aligned}$$

Example 3: Find the volume V of the finite region bounded below by the paraboloid $z = x^2+y^2$, and above by the sphere $x^2+y^2+z^2 = 6$.

This problem clearly has axial symmetry about the z-axis so again do $V = \iiint_V r dr d\phi dz$. First we need to find where the two surfaces meet given in cylindrical coordinates by

$$z = r^2 = \sqrt{6-r^2}$$

which has solution $r = \sqrt{2}$. For each r we need to integrate z in the range

$$r^2 \leq z \leq \sqrt{6-r^2}.$$

The ϕ doesn't really play much role here as it doesn't appear in any of the limits so we can integrate it first. So in total we have

$$\begin{aligned}
V &= \int_0^{\sqrt{2}} \int_{r^2}^{\sqrt{6-r^2}} \int_0^{2\pi} r d\phi dz dr \\
&= 2\pi \int_0^{\sqrt{2}} r \left[z\right]_{r^2}^{\sqrt{6-r^2}} dr \\
&= 2\pi \int_0^{\sqrt{2}} r(\sqrt{6-r^2} - r^2) dr \\
&= 2\pi \left[-\frac{1}{3}(6-r^2)^{\frac{3}{2}} - \frac{r^4}{4} \right]_0^{\sqrt{2}} \\
&= 2\pi(-\frac{1}{3}.8 - \frac{4}{4} + \frac{1}{3}6^{\frac{3}{2}}) \\
&= 2\pi(2\sqrt{6} - \frac{11}{3})
\end{aligned}$$

7.2 Applications: volumes, masses, centres of masses and centroids

Masses and volumes already seen. Centre of mass of a body has coordinates $\bar{x}, \bar{y}, \bar{z}$ where

$$\bar{x} \int \rho \, dV = \int x \, \rho \, dV$$

and similar for y, z.

Example Find the cenre of mass of example 3 above, assuming constant density.

Clearly $\bar{x} = \bar{y} = 0$ by symmetry. Compute $\bar{z}$. Using above formula we may set $\rho = 1$:

$$
\begin{aligned}
\bar{z}V &= \int_0^{\sqrt{2}} \int_{r^2}^{\sqrt{6-r^2}} \int_0^{2\pi} zrd\phi dzdr \\
&= 2\pi \int_0^{\sqrt{2}} r\left[z^2/2\right]_{r^2}^{\sqrt{6-r^2}} dr \\
&= \pi \int_0^{\sqrt{2}} r(6-r^2-r^4)dr \\
&= \pi \left[3r^2 - \frac{r^4}{4} - \frac{r^6}{6}\right]_0^{\sqrt{2}} \\
&= \pi(6-1-\frac{8}{6}) \\
&= \frac{11}{3}\pi
\end{aligned}
$$

giving $\bar{z} = 11\pi/(3V) = 11/(6(2\sqrt{6} - \frac{11}{3})) \approx 1.488$

Definition: Centroid= what the centre of mass of an object would be if it had constant density (even if it doesn't.) (Use above formula with $\rho = 1$)

7.3 Change of variables for triple integrals: Spherical polar coordinates

General change of variable formula for triple integrations is similar to the double integration case. Recall for double integration when changing variables, the measure changed via a Jacobian

$$dA = dxdy = \frac{\partial(x,y)}{\partial(u,v)}dudv = |x_u y_v - x_v y_u|dudv,$$

Now note that the Jacobian can be written as a determinant

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix}.$$

In the same way for triple integration we can change variables and need to multiply the measure by a Jacobian which is the determinant of a $3 \times$ 3determinant. If our new variables are u, v, w then

$$dV = dx\,dy\,dz = \frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{vmatrix}.$$

The proof of this is similar to that for the double integration case. Consider a small volume element with sides $\Delta u, \Delta v, \Delta w$. It is a parallelepiped, with edges given by the 3 vectors

$$\frac{\partial x}{\partial u}\mathbf{i} + \frac{\partial y}{\partial u}\mathbf{j} + \frac{\partial z}{\partial u}\mathbf{k}, \qquad \frac{\partial x}{\partial v}\mathbf{i} + \frac{\partial y}{\partial v}\mathbf{j} + \frac{\partial z}{\partial v}\mathbf{k} \qquad \frac{\partial x}{\partial w}\mathbf{i} + \frac{\partial y}{\partial w}\mathbf{j} + \frac{\partial z}{\partial w}\mathbf{k}$$

The volume of such a parallelepiped is given by the scalar triple product $|\mathbf{a}.(\mathbf{b} \times \mathbf{c})|$ also equal to the determinant above.

7.4 Cylindrical polar coordinates

Here we have coordinates $(r, \phi, z)$where $x = r\cos\phi, y = r\sin\phi, z = z$ so the Jacobian becomes

$$\frac{\partial(x,y,z)}{\partial(r,\phi,z)} = \begin{vmatrix} \cos\phi & \sin\phi & 0 \\ -r\sin\phi & r\cos\phi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r$$

as we saw already.

7.5 Spherical polar coordinates

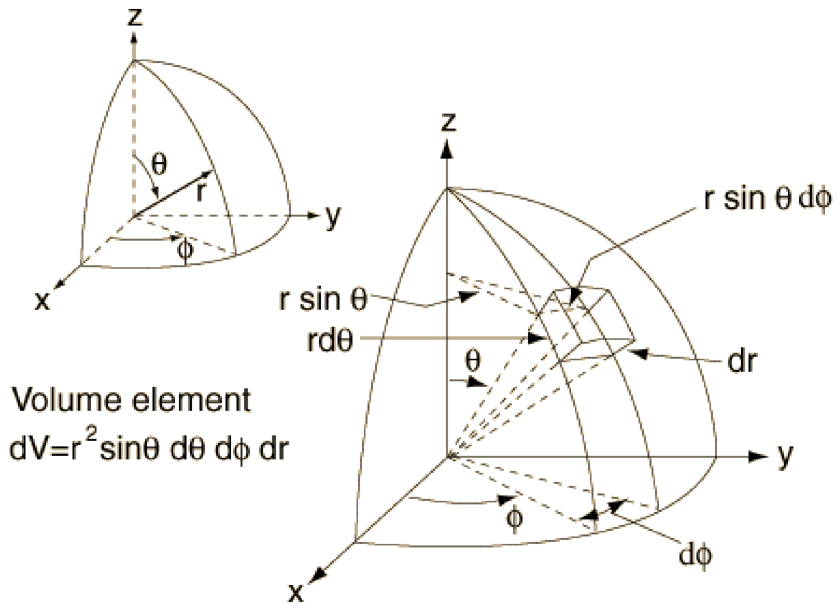
Spherical polar coordinates are useful in cases with radial symmetry, involving for example central charges, gravitating stars, black holes, bubbles, explosions, etc etc. In this case the volume is mapped out by radius r and the two angles, one altidudinal angle θ giving "latitude" and the other "azimuthal" angle ϕ giving "longitude". The coordinates x, y, z are given by

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta. \end{aligned}$$

The Jacobian is then

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = \begin{vmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ r\cos\theta\cos\phi & r\cos\theta\sin\phi & -r\sin\theta \\ -r\sin\theta\sin\phi & r\sin\theta\cos\phi & 0 \end{vmatrix} = r^2\sin\theta$$

It is illustrated in the figure, (to tie in with convention I'll now call the angular coordinate ϕ)



By convention (in maths) the angle θ is taken to be zero at the zenith (i.e. straight up) and π at the nadir (i.e. straight down). If we wish to integrate over the volume the volume elements are curvilinear boxes having sides Δr, $r\sin\theta\Delta\phi$ and $r\Delta\theta$ so that

$$\iiint_V f(x,y,z)\,dxdydz \;=\; \iiint_V f(r\sin\theta\cos\phi, r\sin\theta\sin\phi, z\cos\theta)\,r^2\sin\theta\,drd\phi d\theta$$

Example 4: Find the volume of the region bounded above by a sphere of radius a and below by a cone of half angle $m < \frac{\pi}{2}$.

For this example we may use SP coordinates if we place the axis of the cone along the $\theta = 0$ direction. Then the integration region is

$$\begin{aligned} 0 &\le \phi \le 2\pi \\ 0 &\le \theta \le m \\ 0 &\le r \le a \end{aligned}$$

and

$$
\begin{aligned}
V &= \int_0^a \int_0^m \int_0^{2\pi} r^2 \sin\theta \, d\phi d\theta dr \\
&= 2\pi \int_0^a r^2[-\cos\theta]_0^m \, dr \\
&= 2\pi(1-\cos m) \int_0^a r^2 dr \\
&= \frac{2\pi a^3}{3}(1-\cos m)
\end{aligned}
$$

Note that when $m = \pi$ we get the volume of the sphere $(4\pi a^3)$.

Example 5: Evaluate $I = \iiint_V \frac{\sin(x^2+y^2+z^2)}{\sqrt{x^2+y^2}} dxdydz$ where V is that part of the interior of the sphere $x^2+y^2+z^2 = 9$ in which $x \geq 0$, $y \geq 0$ and $z \geq 0$. i.e. the "positive octant".

The positive octant is given by the integration region is

$$
\begin{aligned}
0 &\leq \phi \leq \frac{\pi}{2} \\
0 &\leq \theta \leq \frac{\pi}{2} \\
0 &\leq r \leq 3
\end{aligned}
$$

Also we have $x^2+y^2 = r^2\sin^2\theta(\sin^2\phi + \cos^2\phi) = r^2\sin^2\theta$, so that

$$
\begin{aligned}
I &= \int_0^3 \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \frac{\sin(r^2)}{r\sin\theta} r^2 \sin\theta \, d\phi d\theta dr \\
&= \int_0^3 \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} r\sin(r^2) \, d\phi d\theta dr \\
&= \frac{\pi^2}{4} \int_0^3 r\sin(r^2) \, dr \\
&= \frac{\pi^2}{4}[-\frac{1}{2}\cos(r^2)]_0^3 \\
&= \frac{\pi^2}{8}(1-\cos(9))
\end{aligned}
$$

8 Vector Calculus

Combine together what we know about vectors with what we know about calculus. Already seen this a little bit...

8.1 Revision of 1st term: vectorial functions of one variable

In A-level: scalar functions of one variable. ODEs: Physical picture, particle moving in 1 dimension $x(t)$. Vectorial function of 1 variable. Picture: Particle traveling in 3d $\mathbf{r}(\mathbf{t})$

8.1.1 Basic formula: (here the variable is t)

$$\frac{d\mathbf{a}(t)}{dt} = \frac{da_1}{dt}\mathbf{i} + \frac{da_2}{dt}\mathbf{j} + \frac{da_3}{dt}\mathbf{k}$$

8.1.2 Leibnitz rules: (See term 1 handout)

8.1.3 Chain rule:

$$\frac{d\mathbf{a}(s)}{du} = \frac{ds}{du}\frac{d\mathbf{a}}{ds}$$

8.1.4 Non-constant basis, eg Polar coordinates: (See term 1 handout)

Often even the basis vectors need to be differentiated. Cartesians $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are constant, but in Polars, for example, they are not constant. Quick recap.

Consider motion in a plane, using polar coordinates r, θ, where $x = r\cos\theta$ and $y = r\sin\theta$. $\mathbf{r} = x\mathbf{i} + y\mathbf{j} = r(\cos\theta\mathbf{i} + \sin\theta\mathbf{j})$.

The **radial unit vector** $\mathbf{e}_r$ is a vector in the direction of $\mathbf{r}$, $\mathbf{e}_r = \frac{\mathbf{r}}{r} = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}$.

The **tangential unit vector** $\mathbf{e}_\theta$ is a vector perpendicular to $\mathbf{e}_r$, and is $\mathbf{e}_\theta = -\sin\theta\mathbf{i} + \cos\theta\mathbf{j}$, (increasing θ is anti-clockwise).

If the particle is moving then r and θ can depend on time. $\dot{\mathbf{e}}_r = \frac{d\theta}{dt}\frac{d}{d\theta}\mathbf{e}_r = \dot{\theta}(-\sin\theta\mathbf{i} + \cos\theta\mathbf{j}) = \dot{\theta}\mathbf{e}_\theta$. $\dot{\mathbf{e}}_\theta = \frac{d\theta}{dt}\frac{d}{d\theta}\mathbf{e}_\theta = \dot{\theta}(-\cos\theta\mathbf{i} - \sin\theta\mathbf{j}) = -\dot{\theta}\mathbf{e}_r$.
Note: $\mathbf{e}_r \cdot \mathbf{e}_r = \mathbf{e}_\theta \cdot \mathbf{e}_\theta = 1$ and $\mathbf{e}_r \cdot \mathbf{e}_\theta = 0$ for all time.

$\mathbf{r} = r\mathbf{e}_r$ therefore $\dot{\mathbf{r}} = \dot{r}\mathbf{e}_r + r\dot{\mathbf{e}}_r = \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta$. $\ddot{\mathbf{r}} = \ddot{r}\mathbf{e}_r + \dot{r}\dot{\mathbf{e}}_r + \dot{r}\dot{\theta}\mathbf{e}_\theta + r\ddot{\theta}\mathbf{e}_\theta + r\dot{\theta}\dot{\mathbf{e}}_\theta = (\ddot{r} - r\dot{\theta}^2)\mathbf{e}_r + (2\dot{r}\dot{\theta} + r\ddot{\theta})\mathbf{e}_\theta$.

8.2 Integration of vectors

(Indefinite) Integration of vectors (or expressions involving vectors) wrt scalar = inverse of differentiation. But rememember: if the expression we integrate is a vector, then

- the integral is also a vector
- constant of integration is also a vector

eg if $\mathbf{a}(t) = d\mathbf{A}/dt$, then the indefinite integral of $\mathbf{A}$is

$$\int \mathbf{a}(t)\,dt = \mathbf{A}(t) + \mathbf{b}$$

where $\mathbf{b}$is a constant vector (since $\mathbf{a}(t)$ is a vector.) As you would expect the definite integral from $t = t_1$to $t = t_2$is

$$\int_{t_1}^{t_2} \mathbf{a}(t)\,dt = \mathbf{A}(t_1) - \mathbf{A}(t_2)$$

In fact you have already seen this. Last term you proved that for a central force, Torque $d\mathbf{L}/dt = 0$ you then concluded that $\mathbf{L}$ =constant. This is by integrating wrt to t.

Example: Integrate $\mathbf{a}(t) = \cos t\,\mathbf{i} + 2t\sin t^2\,\mathbf{j} + 4t^3\,\mathbf{k}$. Integrate term by term so

$$\int \mathbf{a}(t)dt = \sin t\,\mathbf{i} - \cos t^2\,\mathbf{j} + t^4\,\mathbf{k}$$

8.3 Vector functions of several arguments

$\mathbf{a}(u_1, u_2, \ldots, u_n)$vector function of n variables.

8.3.1 Basic formula still applies

$$\frac{\partial \mathbf{a}}{\partial u_r} = \frac{\partial a_1}{\partial u_r}\mathbf{i} + \frac{\partial a_2}{\partial u_r}\mathbf{j} + \frac{\partial a_3}{\partial u_r}\mathbf{k}$$

8.3.2 Chain rule generalises to the multi-variable case

If $\mathbf{a}(u_1, u_2, \ldots, u_n)$ and each $u_i(v_1, v_2, \ldots, v_n)$, then just as for scalar functions we have

$$\frac{\partial \mathbf{a}}{\partial v_i} = \frac{\partial \mathbf{a}}{\partial u_1}\frac{\partial u_1}{\partial v_i} + \frac{\partial \mathbf{a}}{\partial u_2}\frac{\partial u_2}{\partial v_i} + \cdots + \frac{\partial \mathbf{a}}{\partial u_n}\frac{\partial u_n}{\partial v_i}$$

8.3.3 Total differential generalises straightforwardly

If we have a vector function of n variables, $\mathbf{a}(u_1, u_2, \ldots, u_n)$ then the total derivative is:

$$d\mathbf{a} = \frac{\partial \mathbf{a}}{\partial u_1}du_1 + \frac{\partial \mathbf{a}}{\partial u_2}du_2 + \cdots + \frac{\partial \mathbf{a}}{\partial u_n}du_n.$$

8.4 Scalar and Vector fields

Up until now we have considered functions of many variables, often without specifying what the variables are (often space, sometimes space and time, often we didn't care). From now on our variables will **always** be coordinates in space.

A scalar field is simply a function $f(x, y, z)$ where x, y, z are co-ordinates of space (we have considered these previously - but have also considered more general functions eg $f(x, t)$ or just where we didn't care what our variables represented). So to every point in space (x, y, z) is associated a number $f(x, y, z)$. **Example:** pressure in a fluid $P(x, y, z)$**.**

A vector field is a vectorial function of x, y, z. So it associates a **vector** $\mathbf{v}(x, y, z)$ to every point in space (x, y, z). **Example:** velocity vector in a fluid (giving the speed and direction of motion of the fluid at each point.)

Note: although these operations are defined for 3 dimensions, we will often specialise to 2d (simply setting the **k**components to zero.)

8.5 Vector Operators: Grad

There are different differential operators which can be applied to scalar and vector fields. They can all be written in terms of the vector operator, ∇ called *del* or *nabla.* It is defined as

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

So if $f(x, y, z)$ is a scalar field (function of x, y, z) then

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

Example: If $f(x, y, z) = x^2y^3z$ what is ∇f?
$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k} = 2xy^3z\mathbf{i} + 3x^2y^2z\mathbf{j} + x^2y^3\mathbf{k}$

8.5.1 Directional derivatives

The partial derivatives give the rates of change of f when we move along the $\mathbf{i}$ direction or the $\mathbf{j}$ direction. How can we find the rate of change in the direction of a unit vector $\hat{\mathbf{n}}$ in some other direction? A clue comes from the derivation of the chain rule earlier:

If $f = f(x, y, z)$, recall (section 2.4) that if we make a small change $\Delta x, \Delta y, \Delta z$ the change in f, Δf is

$$\begin{aligned}
\Delta f &= \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z. \\
&= \left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}\right) \cdot (\Delta x\mathbf{i} + \Delta y\mathbf{j} + \Delta z\mathbf{k})
\end{aligned}$$

(We used this to define the total differential, there we had only two variables but the extension to 3 is straightforward). This is the total change in f if we move Δx in the $\mathbf{i}$ direction, followed by Δy in the $\mathbf{j}$ direction, followed by Δz in the $\mathbf{k}$ direction. Our position goes from $(x, y, z) \to (x+\Delta x,\ y+\Delta y,\ z+\Delta z)$ which we can write as $\mathbf{x} \to \mathbf{x}+\Delta x\mathbf{i}+\Delta y\mathbf{j}+\Delta z\mathbf{k}+$. Now suppose $\Delta x\mathbf{i}+\Delta y\mathbf{j}+\Delta z\mathbf{k}$ has some infinitessimal length Δs i.e. $((\Delta x)^2+(\Delta y)^2+(\Delta z)^2 = (\Delta s)^2)$. Then $\Delta x\mathbf{i} + \Delta y\mathbf{j} + \Delta z\mathbf{k} = \Delta s\hat{\mathbf{n}}$ and then

$$\begin{aligned}
\Delta f &= \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z \\
&= \Delta s\, \hat{\mathbf{n}}.(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k})
\end{aligned}$$

where the dot indicates the scalar dot product as usual. The vector

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

is called the "gradient of f", (or "grad f" or "del f" for short) and the *vector differential operator* $\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$ is called *del, grad* or *nabla.* So the **rate of change along a direction $\hat{\mathbf{n}}$** is in the limit then given by

$$\frac{df}{ds} = \hat{\mathbf{n}}.\nabla f$$

Note that ∇f is a *vector.* Often using ∇ greatly simplifies manipulations; for example consider Taylor series again. In a space of variables $\mathbf{x} = (x_1, ..x_n)$ we can write the expansion of f at $\mathbf{x} = \mathbf{x_0} + \mathbf{h}$ as

$$f(\mathbf{x_0} + \mathbf{h}) = e^{\mathbf{h}.\nabla} f|_{\mathbf{x_0}}$$

The Laplacian becomes $(\nabla.\nabla)f$ often written $\nabla^2 f$ and so on.

8.5.2 What does ∇f tell us about the surface $f(x, y)$?

Consider separating ∇f into the modulus and unit vector;

$$\nabla f = \hat{\mathbf{m}}|\nabla f|$$

Then the gradient in a direction $\hat{\mathbf{n}}$ is given by

$$(\hat{\mathbf{m}}.\hat{\mathbf{n}})|\nabla f| = \cos\theta|\nabla f|$$

where θ is the angle between the two directions. The gradient is therefore a maximum when $\theta = 0$ - in other words ∇f points *in the direction of maximum gradient* (In 2D, thinking of the function as "height" the gradient points "up the slope").

In 2D, the contours or *level curves* of the surface are by definition the directions along which f is constant. This means $\cos\theta = 0$ - in other words the contours are at right angles to ∇f.

In 3D we have level *surfaces.*

Examples of level curves: contours lines on a map represent height above sea level - lines of constant $h(x, y)$; isobars on a weather map represent lines of constant atmospheric pressure (at sea level) $p(x, y)$.

More generally, in 3 (or more) dimensions ∇f points in the direction that f increases the fastest and is perpendicular to the level-*surfaces.*

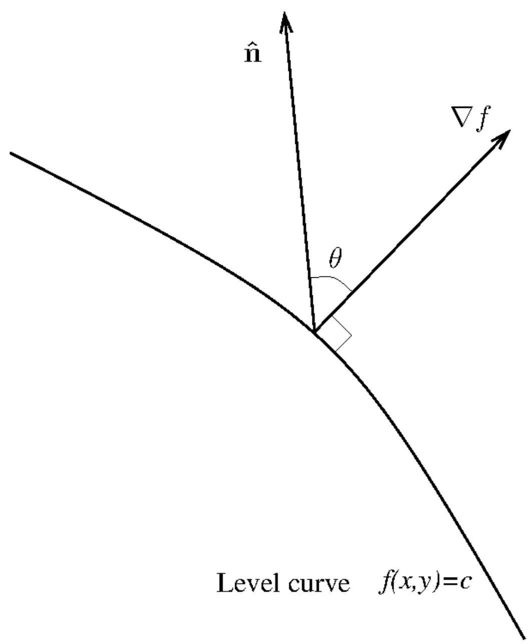


Figure 1: Showing an example level curve given by the equation $f(x, y) = c$ a constant, ∇f at a point (which is a vector perpendicular to the level curve) and an arbitrary unit vector at the same point $\hat{n}$. The rate of change in the direction $\hat{\mathbf{n}}$ (unit vector) is $\hat{\mathbf{n}} \cdot \nabla f = |\nabla f| \cos\theta$ where $\cos\theta$is the angle beween $\hat{\mathbf{n}}$ and ∇f. (as indicated on the figure)

8.5.3 Summary of Grad

- In Cartesian coordinates: $\nabla f = (\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k})$
- Takes a scalar field $f(x, y, z)$ to a vector field ∇f
- ∇f points in the direction of maximum gradient (Grad is short for the "gradient operator".)
- Directional derivative: The rate of change in the direction $\hat{\mathbf{n}}$ (unit vector) is $\hat{\mathbf{n}} \cdot \nabla f$
- Level curves (in 2d - or level surfaces in 3d) is the curves $f(x, y, z) = c$ are perpendicular to ∇f

8.5.4 Examples

Example 1: Find the rate of change of $f(x,y) = y^4 + x^2y^2 + x$ at (0,1) in the direction of the vector $\mathbf{i} + 2\mathbf{j}$.

First find ∇f. We have

$$\begin{aligned} f_x &= 1 + 2xy^2 = 1 \\ f_y &= 4y^3 + 2yx^2 = 4 \\ \nabla f &= (1 + 2xy^2)\mathbf{i} + (4y^3 + 2yx^2)\mathbf{j} = \mathbf{i} + 4\mathbf{j} \end{aligned}$$

Now we need the *unit vector.* This is

$$\begin{aligned} \hat{\mathbf{n}} &= (\mathbf{i} + 2\mathbf{j})/(1 + 2^2) \\ &= \frac{1}{\sqrt{5}}(\mathbf{i} + 2\mathbf{j}) \end{aligned}$$

Then the rate of change is

$$\begin{aligned} \hat{\mathbf{n}}.\nabla f &= \frac{1}{\sqrt{5}}(\mathbf{i} + 2\mathbf{j}).(\mathbf{i} + 4\mathbf{j}) \\ &= \frac{9}{\sqrt{5}} \end{aligned}$$

Example 2: The temperature in a room at the point with coordinates (x, y, z) is given by $T(x, y, z) = x^2e^{-y}z$. At the point (2,1,1) in what direction does the temperature increase most rapidly?

First find ∇T. We have

$$\begin{aligned} T_x &= 2xe^{-y}z = 4/e \\ T_y &= -x^2e^{-y}z = -4/e \\ T_z &= x^2e^{-y} = 4/e \\ \Rightarrow \nabla T &= 2xe^{-y}z\mathbf{i} - x^2e^{-y}z\mathbf{j} + x^2e^{-y}\mathbf{k} = \frac{4}{e}(\mathbf{i} - \mathbf{j} + \mathbf{k}) \end{aligned}$$

So the direction of greatest increase is

$$\frac{1}{\sqrt{3}}(\mathbf{i} - \mathbf{j} + \mathbf{k}).$$

The rate is

$$|\nabla T| = \frac{4\sqrt{3}}{e}\,K/m.$$

Example 3: Find the level curves and gradient of $f(x,y) = x^2 + y^2$.

The level curves are defined by $f(x,y) = c = x^2 + y^2$, so they are circles as expected. Check that this fits with the ∇f. We have $\nabla f = 2(x,y) = 2r(\cos\theta, \sin\theta)$ in polars, which is an arrow of length $2r$ pointing radially from the origin. The unit vector orthogonal points in the direction of the level curve passing through (x,y); $\hat{\mathbf{m}} = \frac{1}{r}(y, -x) = (\sin\theta, -\cos\theta)$.

Example 4: Find the level curve of $f(x,y) = y^4 + x^2y^2 + x$ through the point $(0,1)$ and verify that its tangent at this point is orthogonal to ∇f.

The level curves are defined by $f(x,y) = c = y^4 + x^2y^2 + x$. At (0,1) you can easily verify that $f(0,1) = 1$ so must have $c = 1$. The equation of the level curve is $y^4 + x^2y^2 + x = 1$. The tangent to this point has "slope" $\frac{dy}{dx}$ in the x, y plane. Differentiating we find

$$\begin{aligned}\frac{dy}{dx}(4y^3 + 2yx^2) + 2xy^2 + 1 &= 0 \\ \frac{dy}{dx} &= -\frac{1}{4}.\end{aligned}$$

So the corresponding unit vector in the x, y plane is

$$\hat{\mathbf{p}} = \frac{1}{\sqrt{17}}(4, -1)$$

Next ∇f is given by

$$\nabla f = (2xy^2 + 1,\ 4y^3 + 2yx^2) = (1,\ 4)$$

and is along unit-vector $\hat{\mathbf{m}} = \frac{1}{\sqrt{17}}(1, 4)$. Then we find $\hat{\mathbf{m}}.\hat{\mathbf{p}} = 0$.

Example 5a: Find the rate of change of $f(x,y) = \sin(\sqrt{x^2 + y^2})$ at some position (x,y) and along the direction of the unit vector $\hat{\mathbf{n}} = n_x\mathbf{i} + n_y\mathbf{j}$.

In this and the next example we'll do an example two ways that demonstrates the important concept of *coordinate invariance* of $\hat{\mathbf{n}}.\nabla f$: First we'll use the cartesian coordinates in which the question is given. For ∇f we have

$$\begin{aligned} f_x &= \frac{x\ \cos(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}} \\ f_y &= \frac{y\ \cos(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}} \\ \nabla f &= \frac{\cos(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}\,(x\mathbf{i}+y\mathbf{j}) \end{aligned}$$

so that

$$\begin{aligned} \frac{df}{ds} &= \hat{\mathbf{n}}.\nabla f \\ &= \frac{\cos(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}\,(xn_x+yn_y) \end{aligned}$$

We will repeat this computation in polar coordinates shortly. But first, 2 more vector operators...Div, Curl.

8.6 Vector Operators: Div

The divergence of a vector field $\mathbf{v}(x,y,z)=v_1(x,y,z)\mathbf{i}+v_2(x,y,z)\mathbf{j}+v_3(x,y,z)\mathbf{k}$, "Div $\mathbf{v}$" is defined by

$$\mathrm{div}\,\mathbf{v}=\nabla\cdot\mathbf{v}=\frac{\partial v_1}{\partial x}+\frac{\partial v_2}{\partial y}+\frac{\partial v_3}{\partial z}.$$

This is, of course what you get from $\nabla\cdot\mathbf{v}=(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z})\cdot(v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k})$
Example: Find the divergence of the vector field $\mathbf{v}=x^2y\,\mathbf{i}+xz\cos(y)\,\mathbf{j}+e^{xy}z^3\,\mathbf{k}$.

$$\nabla\cdot\mathbf{v}=2xy-xz\sin(y)+3e^{xy}z^2$$

Note: Div takes a **vector field** and gives a **scalar field.** (The opposite of **Grad**.)
Significance of Div: If we think of the vector field $\mathbf{v}(x,y,z)$ as giving the flow of some quantity, then $\nabla\cdot\mathbf{v}$ gives a measure of the net amount flowing out of any point. (see Wolfram demonstration:
"http://demonstrations.wolfram.com/VectorFieldFlowThroughAndAroundACircle/)

8.6.1 Laplacian: $\nabla^2 f$

We have seen that **Grad** takes a scalar field to a vector field, whereas **Div** takes a vector field to a scalar field. Therefore **Div Grad**, takes a scalar field to a scalar field. What is this operator?

$$\text{div}(\text{grad} f) = \nabla \cdot (\nabla f) = \nabla \cdot (f_x \mathbf{i} + f_y \mathbf{j} + f_z \mathbf{k}) = f_{xx} + f_{yy} + f_{zz}$$

Div(Grad) $=\nabla^2$ is the Laplacian (which we met previously). (Can be defined in any dimension.)

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Example: Find the Laplacian of the scalar field $f(x, y, z) = xy^2 e^{2z}$.

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0 + 2xe^{2z} + 4xy^2 e^{2z} = 2xe^{2z}(1 + 2y^2).$$

8.7 Vector Operators: Curl

The operators **Grad, Div** can be defined for fields in any dimension. The operator **curl** only works in 3 dimensions. It is defined as follows, for any vector field $\mathbf{v}(x, y, z) = v_1(x, y, z)\mathbf{i} + v_2(x, y, z)\mathbf{j} + v_3(x, y, z)\mathbf{k}$:

$$\text{curl}\, \mathbf{v} = \nabla \times \mathbf{v} = \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k}.$$

This is, of course what you get from $\nabla \times \mathbf{v} = (\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}) \times (v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k})$
Example: Find the curl of the vector field $\mathbf{v} = xyz\,\mathbf{i} + yz\,\mathbf{j} + xz\,\mathbf{k}$.

$$\nabla \times \mathbf{v} = -y\,\mathbf{i} + (xy - z)\,\mathbf{j} - xz\,\mathbf{k}$$

Curl gives a measure of the angular velocity of the fluid near a point. (see Wolfram demonstration:
"http://demonstrations.wolfram.com/VectorFieldFlowThroughAndAroundACircle/)

8.7.1 Famous example of Grad, Div, Curl

Maxwell's equations for Electro-magnetism. **E**is the electric field (it is a vector field) **B**is the magnetic field (it is also a vector field). **B**points in the direction of a compass. In a vacuum these are governed by Maxwell's equations which are written in terms of Grad, Div and Curl:

$$\begin{aligned}\nabla \cdot \mathbf{E} =& 0 \\ \nabla \cdot \mathbf{B} =& 0 \\ \nabla \times \mathbf{E} =& -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} =& \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}\end{aligned}$$

More later.....

8.8 Vector operator Formulae [Riley:10.8]

8.8.1 Vector operators (Grad, Div, Curl) acting on Sums/products

1. Grad, Div, Curl are **linear operators** (means "sum then operate=operate then sum") ie:

$$\begin{aligned}\nabla(f+g) =& \nabla f + \nabla g \\ \nabla \cdot (\mathbf{v}+\mathbf{w}) =& \nabla \cdot \mathbf{v} + \nabla \cdot \mathbf{w} \\ \nabla \times (\mathbf{v}+\mathbf{w}) =& \nabla \times \mathbf{v} + \nabla \times \mathbf{w}\end{aligned}$$

   Proof: follows from linearity of $\partial/\partial x$ etc. eg. $\nabla \cdot (\mathbf{v}+\mathbf{w}) = \partial/\partial x(v_1 + w_1) + \partial/\partial y(v_2+w_2) + \partial/\partial z(v_3+w_3) = \partial/\partial x(v_1) + \partial/\partial y(v_2) + \partial/\partial z(v_3) + \partial/\partial x(w_1) + \partial/\partial y(w_2) + \partial/\partial z(w_3) = \nabla \cdot \mathbf{v} + \nabla \cdot \mathbf{w}$

2. When acting on products involving two scalar fields or a scalar field and a vector field we get natural generalisations of the product rule:

$$\begin{aligned}\nabla(fg) =& (\nabla f)g + f(\nabla g) \\ \nabla \cdot (f\mathbf{v}) =& (\nabla f) \cdot \mathbf{v} + f(\nabla \cdot \mathbf{v}) \\ \nabla \times (f\mathbf{v}) =& (\nabla f) \times \mathbf{v} + f(\nabla \times \mathbf{v})\end{aligned}$$

   Proof: All follow in one way or another from the product rule for $\partial/\partial x$etc. Eg to prove the 3rd formula, let us focus on the **k** component only. The $\mathbf{k}$ component of $\nabla \times (f\mathbf{v})$ is:

$$\mathbf{k}\left(\frac{\partial}{\partial x}(fv_2) - \frac{\partial}{\partial y}(fv_1)\right) = \mathbf{k}\left(\frac{\partial f}{\partial x}v_2 - \frac{\partial f}{\partial y}v_1\right) + \mathbf{k}f(\frac{\partial}{\partial x}v_2 - \frac{\partial}{\partial y}v_1)$$

which equals the $\mathbf{k}$component of the RHS. The other components will be similar.
Example: Show that $\nabla \times (f\mathbf{v}) = (\nabla f) \times \mathbf{v} + f(\nabla \times \mathbf{v})$ for $f = x^2y$, $\mathbf{v} = z\mathbf{i} + 2\mathbf{j} + \mathbf{k}$. Then $\nabla \times (f\mathbf{v}) = \nabla \times (x^2yz\mathbf{i} + 2x^2y\mathbf{j} + x^2y\mathbf{k}) = x^2z\mathbf{i} + (x^2y - 2xy)\mathbf{j} + (4xy - x^2z)\mathbf{k}$. On the other hand $\nabla f \times \mathbf{v} = (2xy\mathbf{i} + x^2\mathbf{j}) \times (z\mathbf{i} + 2\mathbf{j} + \mathbf{k}) = x^2z\mathbf{i} - 2xy\mathbf{j} + (4xy - x^2z)\mathbf{k}$ and $f(\nabla \times \mathbf{v}) = x^2y\mathbf{j}$ which sum to the same.

3. Vector operators acting on products of vector fields can be trickier. The first is the simplest as there are only two terms:

$$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = (\nabla \times \mathbf{a}) \cdot \mathbf{b} - \mathbf{a} \cdot (\nabla \times \mathbf{b})$$

[Note the relation with the vector identity (scalar triple product) $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$. In some sense the above identity is a natural combination of the scalar triple product identity and the product rule for differentiation.

Then we also have the following more complicated identities

$$\begin{aligned}\nabla(\mathbf{a} \cdot \mathbf{b}) &= \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) + (\mathbf{a} \cdot \nabla)\mathbf{b} + (\mathbf{b} \cdot \nabla)\mathbf{a} \\ \nabla \times (\mathbf{a} \times \mathbf{b}) &= \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a}) + (\mathbf{b} \cdot \nabla)\mathbf{a} - (\mathbf{a} \cdot \nabla)\mathbf{b}\end{aligned}$$

Again, the second identity is the analogue of the vector triple product identity $\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = (\mathbf{c} \cdot \mathbf{b})\mathbf{a} - (\mathbf{c} \cdot \mathbf{a})\mathbf{b}$ (combined with the product rule.)
Example: Verify $\nabla(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) + (\mathbf{a} \cdot \nabla)\mathbf{b} + (\mathbf{b} \cdot \nabla)\mathbf{a}$ for $\mathbf{a} = x\mathbf{i} + z\mathbf{j}$, $\mathbf{b} = z^2\mathbf{j} + x\mathbf{k}$.
Then $\nabla(\mathbf{a} \cdot \mathbf{b}) = \nabla(z^3) = 3z^2\mathbf{k}$, $\mathbf{a} \times (\nabla \times \mathbf{b}) = (x\mathbf{i} + z\mathbf{j}) \times (-\mathbf{j} - 2z\mathbf{i}) = (2z^2 - x)\mathbf{k}$, $\quad \mathbf{b} \times (\nabla \times \mathbf{a}) = (z^2\mathbf{j} + x\mathbf{k}) \times (-\mathbf{i}) = -x\mathbf{j} + z^2\mathbf{k}$, $\quad (\mathbf{b} \cdot \nabla)\mathbf{a} = (z^2\partial/\partial y + x\partial/\partial z)(x\mathbf{i} + z\mathbf{j}) = x\mathbf{j}$ $\quad (\mathbf{a} \cdot \nabla)\mathbf{b} = (x\partial/\partial x + z\partial/\partial y)(z^2\mathbf{j} + x\mathbf{k}) = x\mathbf{k}$, the result follows.

8.9 Action on the position vector and related functions

We write the position vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and $r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$. Then we have for example

$$\nabla \cdot \mathbf{r} = \frac{\partial}{\partial x}x + \frac{\partial}{\partial y}y + \frac{\partial}{\partial z}z = 3$$

and since $\partial r/\partial x = x/r$ etc. we have

$$\nabla r = (\frac{x}{r})\mathbf{i} + (\frac{y}{r})\mathbf{j} + (\frac{z}{r})\mathbf{k} = \frac{\mathbf{r}}{r} = \mathbf{e}_r.$$

Exercise: Show that $\nabla f(r) = (df/dr)\mathbf{e}_r$
Then we can use this result with the previous result and identity 1b above, to show that for example:

$$\nabla \cdot (g(r)\mathbf{r}) = 3g(r) + r\frac{dg}{dr}$$

Proof: $\nabla \cdot (g(r)\mathbf{r}) = g(r)\nabla \cdot \mathbf{r} + \nabla g \cdot \mathbf{r} = 3g(r) + \frac{dg}{dr}\mathbf{e}_r \cdot \mathbf{r} = 3g(r) + r\frac{dg}{dr}$

We can also prove:

$$\nabla^2 f(r) = \frac{d^2 f}{dr^2} + \frac{2}{r}\frac{df}{dr}$$

Proof: $\nabla^2 f(r) = \nabla \cdot (\nabla f(r)) = \nabla \cdot (\frac{1}{r}\frac{df}{dr}\mathbf{r})$ using the result for ∇f above. Now we can use the identity for $\nabla \cdot (g(r)\mathbf{r})$ above with $g(r) = \frac{1}{r}\frac{df}{dr}$ to get $\nabla \cdot (\frac{1}{r}\frac{df}{dr}\mathbf{r}) = \frac{3}{r}\frac{df}{dr} + r\frac{d}{dr}(\frac{1}{r}\frac{df}{dr}) = \frac{d^2 f}{dr^2} + \frac{2}{r}\frac{df}{dr}$.

8.10 Combinations of Grad, Div and Curl

We can also consider applying more than one of Grad, Div or Curl on a single scalar or vector field. We get two identities that give zero

$$\begin{aligned}
\mathrm{curl}(\mathrm{grad} f) =& \nabla \times (\nabla f) = 0 \\
\mathrm{div}(\mathrm{curl}\,\mathbf{v}) =& \nabla \cdot (\nabla \times \mathbf{v}) = 0
\end{aligned}$$

So the grad of a function ∇f gives a vector field which has zero curl ("irrotational"). Eg $f = xyz^2$. $\nabla f = yz^2\mathbf{i} + xz^2\mathbf{j} + 2xyz\mathbf{k}$. Check $\nabla \times (\nabla f) = \mathbf{i}(2xz - 2xz) + \mathbf{j}(2yz - 2yz) + \mathbf{k}(z^2 - z^2) = \mathbf{0}$.

Also the curl of a vector field is divergenceless.

The other useful identity is

$$\mathrm{curl}(\mathrm{curl}\,\mathbf{v}) = \nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \nabla^2 \mathbf{v}$$

where we recall that $\nabla^2 f = \nabla \cdot \nabla f$ (although here it is applied to a vector field.)

Example: Show that Maxwell's equations in a vacuum imply that $\mathbf{E}$ and $\mathbf{B}$ satisfy the wave equation. Recall

$$\begin{aligned}
\nabla \cdot \mathbf{E} =& 0 \\
\nabla \cdot \mathbf{B} =& 0 \\
\nabla \times \mathbf{E} =& -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \times \mathbf{B} =& \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}
\end{aligned}$$

Take Curl of equation 2. LHS$=\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\nabla^2 \mathbf{E}$ whereas RHS$=-\nabla \times \frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial(\nabla \times \mathbf{B})}{\partial t} = -\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}$. Thus we have the wave equation $\nabla^2 \mathbf{E} = \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}$. We get the same for $\mathbf{B}$. Must be more information as there are 4 equations.

8.11 General Curvilinear Coordinates

Instead of x, y, z it is sometimes useful to use different co-ordinates u_1, u_2, u_3 (eg Spherical or Cylindrical Polar co-ordinates) so that

$$x = x(u_1, u_2, u_3) \qquad y = y(u_1, u_2, u_3) \qquad z = z(u_1, u_2, u_3)$$

8.11.1 Basis vectors

In cartesians the position vector is

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

and **i**,**j**,**k** are basis vectors. Note that the basis vectors can be given as

$$\mathbf{i} = \frac{\partial \mathbf{r}}{\partial x} \qquad \mathbf{j} = \frac{\partial \mathbf{r}}{\partial y} \qquad \mathbf{k} = \frac{\partial \mathbf{r}}{\partial z}.$$

In general co-ordinates it is useful to similarly define basis vectors in terms of partial differentials of the position vector $\frac{\partial \mathbf{r}}{\partial u_i}$. If the co-ordinates are *orthogonal* this means $\frac{\partial \mathbf{r}}{\partial u_i} \cdot \frac{\partial \mathbf{r}}{\partial u_j} = 0$ if $i \neq j$. It is then useful to make the basis vectors *orthonormal* by ensuring they are of unit length, ie dividing by the modulus. So for an orthogonal system of co-ordinates **(which we assume from now on)** we define basis vectors:

$$\mathbf{e_1} = \frac{1}{h_1}\frac{\partial \mathbf{r}}{\partial u_1} \qquad \mathbf{e_2} = \frac{1}{h_2}\frac{\partial \mathbf{r}}{\partial u_2} \qquad \mathbf{e_3} = \frac{1}{h_3}\frac{\partial \mathbf{r}}{\partial u_3}$$

where h_1.... are called *scale factors:*

$$h_1 = \left|\frac{\partial \mathbf{r}}{\partial u_1}\right| \qquad h_2 = \left|\frac{\partial \mathbf{r}}{\partial u_2}\right| \qquad h_3 = \left|\frac{\partial \mathbf{r}}{\partial u_3}\right|$$

Example 1: Cylindrical polars
[Notes not complete here]
Define basis vectors, scale factors
Derive Grad in general orthogonal coordinates.
Give Cylindrical polars and spherical polars as examples.

Example 5b: Find the rate of change of $f(x, y) = \sin(\sqrt{x^2 + y^2})$ at some position (x, y) and along the direction of the unit vector $\hat{\mathbf{n}} = n_x\mathbf{i} + n_y\mathbf{j}$, using polar coordinates. [This example was first done in Cartesian coordinates in section8.5.4]

Seeing that the function is a function of $r = \sqrt{(x^2 + y^2)}$ only (i.e. $f = \sin r$) you might prefer to use polar coordinates (same as cylindrical coordinates, without the z coordinate). Of course the answer should be the same [We got $\frac{df}{ds} = \hat{\mathbf{n}}.\nabla f = \frac{\cos(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}\,(xn_x + yn_y)$before]. The transformation to polar coordinates is

$$x = r\cos\theta \;\; ; \;\; y = r\sin\theta.$$

In polar coordinates $\nabla f = \mathbf{e_r}\frac{\partial f}{\partial r} + \mathbf{e}_\theta \frac{1}{r}\frac{\partial f}{\partial \theta}$ so that ∇f takes a particularly simple form

$$\nabla f = \cos r\,\mathbf{e_r} + 0\mathbf{e}_\theta = \cos r\,\hat{r}$$

so that

$$\hat{\mathbf{n}}.\nabla f = n_r \cos r$$

where in the polar coordinates $\hat{\mathbf{n}} = n_r \mathbf{e_r} + n_\theta \mathbf{e}_\theta$. So all that remains is to compare with the previous answer:

$$\hat{\mathbf{n}}.\nabla f = \frac{\cos(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}\,(xn_x + yn_y) = \frac{\cos(r)}{r}\,(xn_x + yn_y).$$

Now $x/r = \cos\theta$ and $y/r = \sin\theta$. So the previous answer can be written

$$\hat{\mathbf{n}}.\nabla f = \cos(r)\,(\cos(\theta)n_x + \sin(\theta)n_y) = \cos r\,\mathbf{e_r}\cdot\hat{\mathbf{n}}$$

where we recall $\mathbf{e_r} = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}$ and $\hat{\mathbf{n}} = n_x\mathbf{i} + n_y\mathbf{j}.$ But then by definition $\mathbf{e_r}\cdot\hat{\mathbf{n}} = n_r$ and we get the same answer.