Calculus of Two or More Variables

Lecturer: Dr Eamonn Gaffney

Notes extensively based on material written by Dr Richard Earl

Michaelmas Term 2007

0.0 Syllabus

7. Introduction to Partial Derivatives. Standard Co-ordinate Systems.
8. Chain rule. Change of Variable.
9. Jacobians of Two Variable Systems.
10. Calculation of Areas. Standard Curves and Surfaces.
11. Gradient Vector. Directional Derivatives.
12. Normal to a Surface. Identities.
13. Laplacian and its Form in Other Co-ordinate Systems.
14. Elementary PDEs with motivation for where they arise in applications.
15. Brief introduction to Laplace's equation, Poisson's equation, the wave equation and the diffusion equation.
16. Verification of solutions to these equations.

0.1 Recommended Texts and Website

- D. W. Jordan & P. Smith, *Mathematical Techniques*, 3rd Edition, Oxford (2002), Chapters 27-29, 31.
- Erwin Kreyszig, *Advanced Engineering Mathematics*, 8th Edition, Wiley (1999) Appendix 3.2, Sections 8.8, 8.9, 9.3, 9.5, 9.6.
- http://www.maths.ox.ac.uk/~earl/
- http://www.maths.ox.ac.uk/~gaffney/

1. PARTIAL DIFFERENTIATION

We began our study of ordinary differential equations (ODEs) by modelling the flight of a projectile. Let's consider here, before discussing **partial derivatives** and **partial differential equations** (PDEs), a slightly more complex example of a mug of tea cooling down on a table top.

In what sense is this new physical scenario more complicated to model? The quantity we are naturally interested in here is the temperature T of the liquid. In our earlier motivating example of the projectile, its height depended only on time t. Let's take the mug's dimensions as radius R and height H. Here the temperature of the tea will again depend on time t, but almost certainly not in a uniform way throughout the mug. So rather than just depending on time t the temperature will also depend on the spatial co-ordinates x, y and z. Also heat will be lost to the air, down the sides of the mug, to the air beyond and may heat up the table and surrounding air.

The differential equation governing the behaviour of $T(x, y, z, t)$ is a partial differential equation. The **heat equation**, which we will meet again later in more detail, states that

$$\frac{\partial T}{\partial t} = \kappa \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right)$$

where κ is known as thermal diffusivity. These "fancy" derivatives with the curly ∂s just reflect that T depends on several variables.

- The partial derivative $\partial T/\partial t$ is the derivative of T with respect to time t when we keep all the other variables x, y, z constant.
- Again $\partial T/\partial t$ is a function of the four variables x, y, z, t.
- It is a measure of what we'd see if we focus separately on each point (x, y, z) and watch how the temperature changes over time.

Because the tea is cooling we would expect that $\partial T/\partial t < 0$ throughout the mug and at all times.

What other information would we need to describe the tea's cooling? Well, we would again need an **initial condition** saying how hot the tea was to begin with. So we might assume it to begin uniformly hot and have

$$T(x, y, z, 0) = T_0$$

for some initial temperature T_0 and for $x^2 + y^2 < R^2,\ 0 < z < H$.

We would also need to make some assumptions about how the heat dissipated out of the mug. If we assume that air remains constantly at some ambient room temperature T_A then one **boundary condition** would be

$$T(x, y, H, t) = T_A$$

for $x^2 + y^2 < R^2, t > 0$; this describes the temperature's behaviour at the top of the mug (though it isn't a particularly realistic assumption physically!). If say the mug were insulated to allow no heat loss then another boundary condition would be

$$\frac{\partial T}{\partial z}(x, y, 0, t) = 0$$

for $x^2 + y^2 < R^2, t > 0$ at the base of the mug. Likewise down the sides of the mug we would have

$$\frac{\partial T}{\partial r}(x, y, z, t) = 0$$

for $x^2 + y^2 = R^2, 0 < z < H, t > 0$ and where r denotes the distance of a point from the central axis of the mug.

Indeed, having noticed this symmetry in the cooling we might decide to change variables and consider this as a problem in just three variables r, t, z reducing the problem somewhat.

In this course, these are the sort of PDEs we shall consider, initial and boundary conditions, and how changes of variable can simplify them.

1.1 Partial Derivatives

Definition 1 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a function of* n *variables* $x_1, x_2, \ldots, x_n$*. Then the* ***partial derivative***

$$\frac{\partial f}{\partial x_i}\left(p_1, \ldots, p_n\right)$$

is the rate of change of f*, at* $\left(p_1, \ldots, p_n\right)$*, when we vary only the variable* x_i *about* p_i *and keep all of the other variables constant. Precisely then*

$$\frac{\partial f}{\partial x_i}\left(p_1, \ldots, p_n\right) = \lim_{h \to 0} \frac{f\left(p_1, \ldots, p_{i-1}, p_i + h, p_{i+1}, \ldots p_n\right) - f\left(p_1, \ldots, p_n\right)}{h}.$$

By contrast, derivatives such as df/dx *are sometimes referred to as* ***full derivatives***

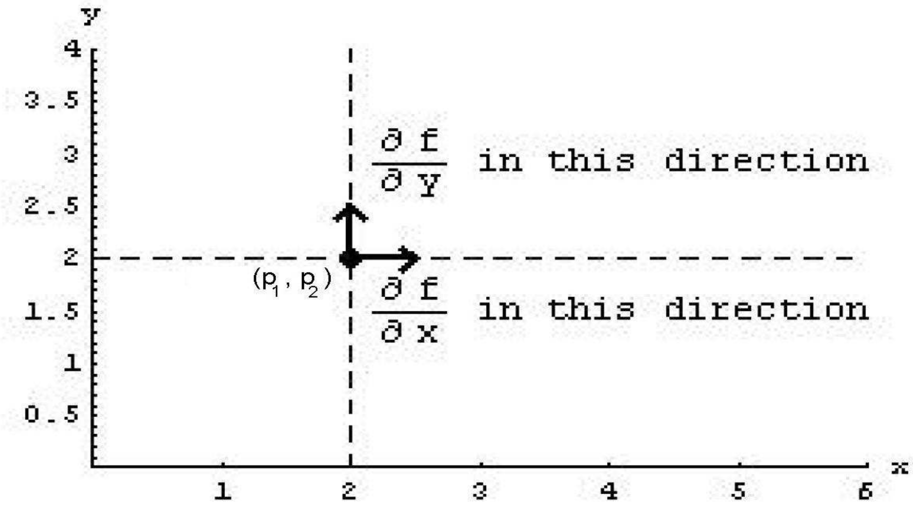
Remark 2 *If* $f(x)$ *is a function of a single variable note that*

$$\frac{df}{dx} = \frac{\partial f}{\partial x}.$$

Notation 3 $\partial f/\partial x$ *is still pronounced "d f by d x" or sometimes "partial d f by d x".*

We shall also, occasionally, write f_x *for* $\partial f/\partial x$*. This is common notation, but as the* $\partial f/\partial x$ *is more visible, we shall mainly stick with that notation through the course.*

In some texts $\partial f/\partial x_i$ *is denoted as* f_i*. It is also sometimes written* $f_{,i}$*.*



$$\frac{\partial f}{\partial x}(a,b)=\lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h}, \qquad \frac{\partial f}{\partial y}(a,b)=\lim_{k\to 0}\frac{f(a,b+k)-f(a,b)}{k}.$$

Notation 4 *To stress which variables are being kept constant some texts use the notation*

$$\left.\frac{\partial f}{\partial x}\right|_y$$

to denote that this is the partial derivative of f with respect to x whilst keeping y constant. This is a measure of how quickly $f(x,y)$ is changing as we move from the point (p_1,p_2) along the line $y=p_2$.

Unless such a notation is used, we will always make clear which ***co-ordinate system*** *we are using and then a partial derivative will always denote differentiation with respect to some co-ordinate with all other co-ordinates in the system being kept constant.*

Example 5 *Let*

$$f(x,y,z) = x^2 + ye^x + \frac{z}{y}.$$

Then

$$\frac{\partial f}{\partial x} = \qquad\qquad\qquad , \qquad \frac{\partial f}{\partial y} = \qquad\qquad\qquad , \qquad \frac{\partial f}{\partial z} = \qquad\qquad\qquad .$$

Example 6 *Let*

$$f(x_1, x_2, x_3, x_4) = \sin(x_1 x_2) + \cos x_4.$$

Then

$$\frac{\partial f}{\partial x_1} = x_2 \cos(x_1 x_2), \qquad \frac{\partial f}{\partial x_2} = x_1 \cos(x_1 x_2), \qquad \frac{\partial f}{\partial x_3} = 0, \qquad \frac{\partial f}{\partial x_4} = -\sin x_4.$$

Note that $\partial f / \partial x_3 = 0$ *as the definition of* f *is independent of* x_3.

Example 7 *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *be defined by*

$$f(x,y) = \begin{cases} x + y & \text{if } x = 0 \text{ or } y = 0, \\ 1 & \text{otherwise} \end{cases}$$

Note that the partial derivatives $\partial f / \partial x$ *and* $\partial f / \partial y$ *exist at* $(0,0)$ *with*

$$\frac{\partial f}{\partial x}(0,0) = 1 \quad \text{and} \quad \frac{\partial f}{\partial y}(0,0) = 1$$

yet f *is not continuous at* $(0,0)$. *This contrasts with full derivatives where differentiability implies continuity (see Hilary term analysis course).*

Definition 8 *We may define second and higher partial derivatives in a similar manner to how we define them for full derivatives. So, in the case of second partial derivatives of a function* $f(x,y)$

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), &\quad \text{also written } f_{xx} = (f_x)_x\,, \\
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), &\quad \text{also written } f_{yx} = (f_y)_x\,, \\
\frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), &\quad \text{also written } f_{xy} = (f_x)_y\,, \\
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right), &\quad \text{also written } f_{yy} = (f_y)_y\,.
\end{aligned}$$

Example 9 *Let us return to the function*

$$f(x,y,z) = x^2 + ye^x + \frac{z}{y}$$

from Example 5. Then

$$\frac{\partial f}{\partial x} = 2x + ye^x, \qquad \frac{\partial f}{\partial y} = e^x - \frac{z}{y^2}, \qquad \frac{\partial f}{\partial z} = \frac{1}{y}.$$

So

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= 2 + ye^x, & \frac{\partial^2 f}{\partial y \partial x} &= e^x, & \frac{\partial^2 f}{\partial z \partial x} &= 0, \\
\frac{\partial^2 f}{\partial x \partial y} &= e^x, & \frac{\partial^2 f}{\partial y^2} &= \frac{2z}{y^3}, & \frac{\partial^2 f}{\partial z \partial y} &= \frac{-1}{y^2}, \\
\frac{\partial^2 f}{\partial x \partial z} &= 0, & \frac{\partial^2 f}{\partial y \partial z} &= \frac{-1}{y^2}, & \frac{\partial^2 f}{\partial z^2} &= 0.
\end{aligned}$$

Note in the previous example

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}, \qquad \frac{\partial^2 f}{\partial z \partial x} = \frac{\partial^2 f}{\partial x \partial z}, \qquad \frac{\partial^2 f}{\partial z \partial y} = \frac{\partial^2 f}{\partial y \partial z}.$$

This will typically be the case in the examples we'll see, as the following theorem shows:

Theorem 10 *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be such that $\frac{\partial^2 f}{\partial y \partial x}$ and $\frac{\partial^2 f}{\partial x \partial y}$ exist and are continuous. Then*

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}.$$

Proof. (For reference only — requires Hilary term analysis, and non-examinable). For $x, y, h, k \in \mathbb{R}$ define

$$\phi(x, y) = f(x, y + k) - f(x, y), \quad \text{and} \quad \psi(x, y) = f(x + h, y) - f(x, y)$$

so that

$$\begin{aligned} D(x, y) &= f(x + h, y + k) - f(x + h, y) - f(x + h, y) + f(x, y) \\ &= \phi(x + h, y) - \phi(x, y) \\ &= \psi(x, y + k) - \psi(x, y). \end{aligned}$$

By the Mean-Value Theorem, applied twice, there exist $\theta_1, \theta_2 \in (0, 1)$ such that

$$\begin{aligned} D(x, y) &= \phi(x + h, y) - \phi(x, y) = h\phi_x(x + \theta_1 h, y) \\ &= h[f_x(x + \theta_1 h, y + k) - f_x(x + \theta_1 h, y)] \\ &= hkf_{xy}(x + \theta_1 h, y + \theta_2 k). \end{aligned}$$

Arguing similarly with the $D(x, y) = \psi(x, y + k) - \psi(x, y)$ expression we know there exist $\theta_3, \theta_4 \in (0, 1)$ such that

$$D(x, y) = hkf_{yx}(x + \theta_3 h, y + \theta_4 k).$$

So

$$f_{xy}(x + \theta_1 h, y + \theta_2 k) = f_{yx}(x + \theta_3 h, y + \theta_4 k)$$

Letting h and k tend to 0, and using the continuity of f_{xy} and f_{yx} we see that $f_{xy} = f_{yx}$ as required. ■

Example 11 *Define* $f : \mathbb{R}^2 \to \mathbb{R}$ *by*

$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Show that

$$\frac{\partial^2 f}{\partial y \partial x} \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y}$$

are unequal at $(0,0)$.

Solution. We have for $x \neq 0$ and $y \neq 0$,

$$\frac{\partial f}{\partial x}(0,y) =$$

and similarly

$$\frac{\partial f}{\partial y}(x,0) =$$

Hence

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1 \neq 1 = \frac{\partial^2 f}{\partial x \partial y}(0,0).$$

■

Example 12 *Find all solutions of the form $f(x, y)$ to the partial differential equations*

$$(i) \qquad \frac{\partial^2 f}{\partial y \partial x} = 0, \qquad (ii) \qquad \frac{\partial^2 f}{\partial x^2} = 0.$$

Solution. The first PDE is

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 0.$$

Those functions $g(x, y)$ which satisfy $\partial g/\partial y = 0$ are functions $p(x)$ which solely depend on x. So we have

$$\frac{\partial f}{\partial x} = p(x).$$

This looks like an equation we would normally just integrate up, not forgetting a constant. But again $\partial/\partial x$ sends to zero any function $Q(y)$ of y. So instead of a constant we have an arbitrary function of y. The solution then is

$$f(x, y) = P(x) + Q(y)$$

where $P(x)$ is an anti-derivative of $p(x)$, i.e. $P'(x) = p(x)$.

For the second equation $\partial^2 f/\partial x^2 = 0$ we can integrate in similar fashion to get

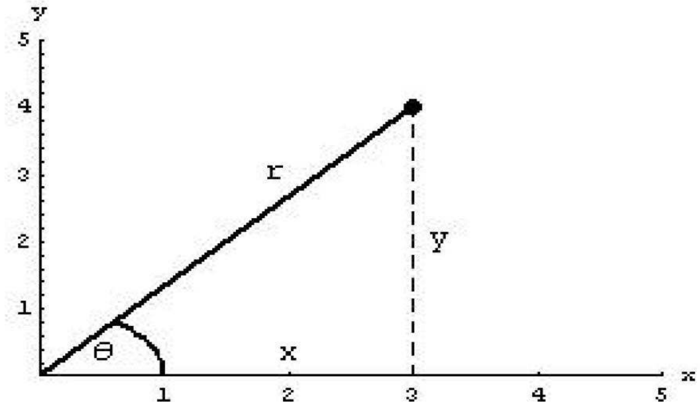
Remark 13 *Note how the solutions to these second order PDEs include two arbitrary* ***functions*** *rather than two arbitrary* ***constants*** *as is often the case of an ODE. This makes sense though when we note that partially differentiating wrt x annihilates functions solely in the variable y and not just constant functions.*

1.2 Co-ordinate Systems

In the examples we have seen so far we have considered functions $f(x,y)$ or $g(x,y,z)$ where we have been thinking of x,y,z as Cartesian co-ordinates. f and g are then functions defined on a 2-dimensional plane or in 3-dimensional space. There are other natural ways to place co-ordinates on a plane or space. Indeed, depending on the nature of a problem and its inherent symmetry, it may be very natural to use other co-ordinates.

In the next lecture we will derive the *chain rule* which describes the relationship between derivatives in one system and another. For now, we simply introduce some important examples of co-ordinate systems.

Example 14 ***(Planar Polar Co-ordinates)*** *Instead of considering a point in a plane as given distances along each of two perpendicular axes, we can equally describe a point P as being at a certain distance r from an origin O and such that OP makes an anti-clockwise angle θ with a fixed axis, usually the positive x-axis. This is shown in the diagram below.*



We can see that the equations relating ***Cartesian co-ordinates*** *and* ***planar polar co-ordinates*** *are*

$$
\begin{aligned}
x &= r\cos\theta, \qquad y = r\sin\theta \\
r &= \sqrt{x^2+y^2}, \qquad \tan\theta = y/x.
\end{aligned}
$$

So in the above diagram where $x = 3$ and $y = 4$ then

$$r = \sqrt{3^2 + 4^2} = 5 \quad \textit{and} \quad \theta = \tan^{-1}\frac{4}{3} \approx 1.249 \ \textit{radians.}$$

Note that r takes values in the range $r \in [0, \infty)$ and $\theta \in [0, 2\pi)$,for example, or equally $(-\pi, \pi]$. Note also that θ is undefined at the origin.

Remark 15 *Note the definition of partial derivative $\partial f/\partial x$ very much depends on the co-ordinate system that x is part of. It is important to know which other co-ordinates are being kept fixed.*

For example, we could have two different co-ordinate systems, one the standard Cartesian co-ordinates and the other being x and the polar co-ordinate θ. Consider now what $\partial r/\partial x$ means in each system.

In Cartesian co-ordinates, we have

$$r = \sqrt{x^2 + y^2} \quad \textit{and so} \quad \partial r/\partial x =$$

However when we write r in terms of x and θ we have

$$r = x/\cos\theta \quad \textit{and so} \quad \frac{\partial r}{\partial x} =$$

which is certainly a different answer! The reason is that the two derivatives we have calculated are

$$\left.\frac{\partial r}{\partial x}\right|_y \quad \textit{and} \quad \left.\frac{\partial r}{\partial x}\right|_\theta$$

and so are measuring the change in x along curves $y =$ const. or along $\theta =$ const. which are very different directions.

Indeed note that the two are equal only when $\cos^2\theta = 1$ in which case lines of constant θ and constant y are in the same direction.

Example 16 ***(Changing from polar to Cartesian co-ordinates and vice versa)***

Given a curve with equation $f(x,y)=0$ *then it can be rephrased as an equation in polar co-ordinates as*

$$g(r,\theta)=f(r\cos\theta, r\sin\theta)=0.$$

The unit circle $x^2+y^2=1$ *clearly becomes* $r=1$ *and the line* $x=k$ *becomes* $r=k\sec\theta$.

More generally, the line $ax+by+c=0$ *becomes* $r(a\cos\theta+b\sin\theta)=-c$ *which can be rewritten as*

$$r=$$

where $A=-c/\sqrt{a^2+b^2}$ *and* $\tan\alpha=b/a$.

In reverse, given a curve in polar co-ordinates $F(r,\theta)=0$, *then this can be rewritten as*

$$G(x,y)=F\left(\sqrt{x^2+y^2},\tan^{-1}\left(\frac{y}{x}\right)\right)=0.$$

For example, $r=\cos\theta$ *becomes* $r^2=r\cos\theta$ *which gives* $x^2+y^2=x$. *This we see is the circle*

$$\left(x-\frac{1}{2}\right)^2+y^2=\left(\frac{1}{2}\right)^2$$

which is a circle of radius $1/2$ *around the point* $(1/2,0)$.

Example 17 ***(Planar Parabolic Co-ordinates)***

Planar Parabolic co-ordinates u, v are given in terms of x and y by the relations

$$x = \frac{1}{2}\left(u^2 - v^2\right), \qquad y = uv.$$

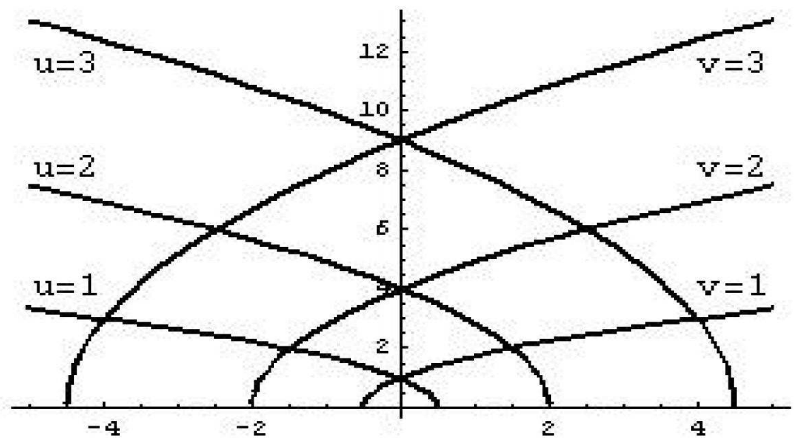
Note that the curve $u = c$, in Cartesian co-ordinates, is

$$2xc^2 = c^4 - y^2$$

and the curve $v = k$, in Cartesian co-ordinates is

$$2xk^2 = y^2 - k^4$$

both of which are parabolas. As u and v vary over the positive numbers then (x, y) varies over the upper half-plane.



Example 18 ***(Cylindrical Polar Co-ordinates)***

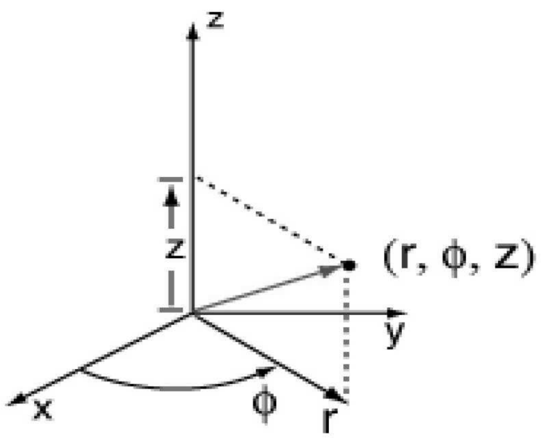
We can naturally extend planar polar co-ordinates into three dimensions using a further z co-ordinate. In this case they are called cylindrical polar co-ordinates. The relationships between r, ϕ, z and x, y, z are given by

$$x = r\cos\phi, \qquad y = r\sin\phi, \qquad z = z$$

and

$$r = \sqrt{x^2 + y^2}, \qquad \tan\phi = \frac{y}{x}, \qquad z = z.$$

Note that $r =$ const. defines a cylinder, $\phi =$ const. defines a vertical plane through the origin and $z =$ const. defines a horizontal plane.



Example 19 *(Spherical Polar Co-ordinates)*

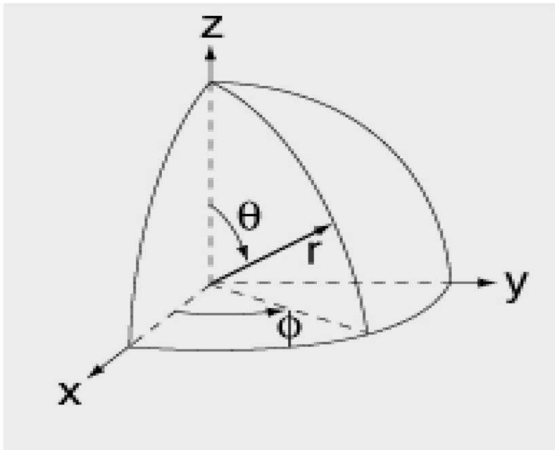
In the same way that latitude and longitude are used to determine a position on the planet, we can similarly use two angles θ and ϕ to determine position on concentric spheres distance r from the origin. These are called spherical polar co-ordinates.

Here r is simply the distance of a point P from the origin O. The angle θ is the angle OP makes with the vertical z-axis and takes values in the range $-\pi/2 \leqslant \theta \leqslant \pi/2$. Finally, if Q is the projection of P vertically into the xy-plane then ϕ is the angle OQ makes with the positive x-axis. The relationships between x, y, z and r, ϕ, θ are given by

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta,$$

and

$$r = \sqrt{x^2 + y^2 + z^2}, \qquad \tan\phi = \frac{y}{x}, \qquad \tan\theta = \frac{\sqrt{x^2+y^2}}{z}.$$

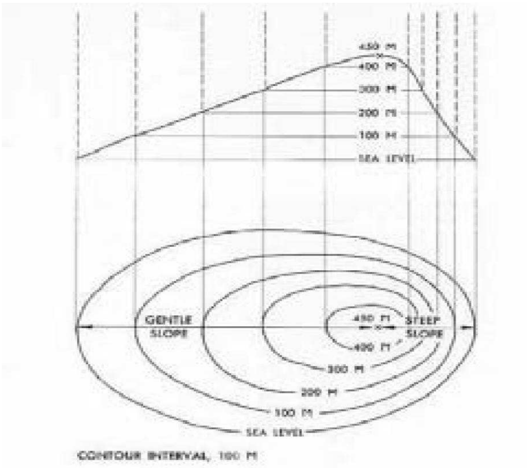


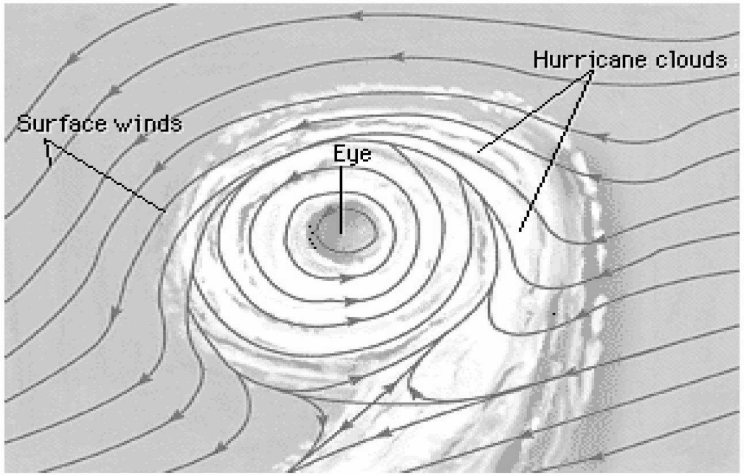
2. CHANGE OF VARIABLE. CHAIN RULE

2.1 Functions of Two Variables

Functions of more than one variable are common throughout mathematics. The motivating example for partial derivatives at the start of the last chapter involved a temperature T which depended on three spatial co-ordinates x, y, z and one temporal co-ordinate t.

Such functions are often associated with maps as well. For example, we might have a physical map denoting the height z above a point (x, y). The map might also include contours which are the curves $z = c$ of constant height. Here z is a *scalar* function of two variables.





Or the map might be a meteorological map denoting the wind speed and direction (at a fixed height) above a point (x, y). Below is a wind-direction field associated with a hurricane — with each point (x, y) is associated a *vector* $\mathbf{u}(x, y)$. That is, $\mathbf{u}$ is a vector-valued function of two variables.

Example 20 *The height z of a column above the point (x, y) is given by the function*

$$f(x, y) = 10 - (x-1)^2 + y^2.$$

Write this as a function $g(r, \theta)$ of planar polar co-ordinates.

Solution. As $x = r\cos\theta$ and $y = r\sin\theta$ then we can write

$$\begin{aligned} f(x, y) &= 10 - (r\cos\theta - 1)^2 + (r\sin\theta)^2 \\ &= 10 - r^2\cos^2\theta + 2r\cos\theta - 1 + r^2\sin^2\theta \\ &= 9 - r^2\cos 2\theta + 2r\cos\theta \\ &= g(r, \theta). \end{aligned}$$

■

Remark 21 *Note that f and g are different functions, even though they have the same numerical value $z = f(x, y) = g(r, \theta)$. It is* ***not*** *the case that $z = f(r, \theta)$, rather z is given by a different rule in terms of r and θ. The rule*

$$z = f(r, \theta) = 10 - (r-1)^2 + \theta^2$$

is clearly not the right one!

2.2 The Chain Rule

The *chain rule for two or more variables* serves the same purpose as the chain rule for one variable which states that

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}.$$

The rule arises when we wish to calculate the derivative of the composition of two functions $f\left(u\left(x\right)\right)$ with respect to x.

Likewise we might have a function $f\left(u,v\right)$ of two variables u and v, each of which are functions of variables x and y. We can then make the composition F

$$F\left(x,y\right) = f\left(u\left(x,y\right),v\left(x,y\right)\right),$$

which is itself a function of x and y. We might then wish to calculate its partial derivatives

$$\frac{\partial F}{\partial x} \quad \text{and} \quad \frac{\partial F}{\partial y}.$$

The chain rule states that

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

Before going on to prove the chain rule, here is an example approached two different ways.

Example 22 *Let*

$$f(u,v) = (u-v)\sin u + e^v, \qquad u(x,y) = x^2 + y, \qquad v(x) = y - 2x,$$

and let $F(x,y) = f(u(x,y), v(x,y))$. *Calculate* $\partial F/\partial x$ *and* $\partial F/\partial y$ *by (i) direct calculation, (ii) the chain rule.*

Solution. (i) We have that

$$F(x,y) = \left(x^2+2x\right)\sin\left(x^2+y\right) + \exp(y-2x).$$

Hence

$$\begin{aligned} \frac{\partial F}{\partial x} &= (2x+2)\sin\left(x^2+y\right) + 2x\left(x^2+2x\right)\cos\left(x^2+y\right) - 2\exp(y-2x); \\ \frac{\partial F}{\partial y} &= \left(x^2+2x\right)\cos\left(x^2+y\right) + \exp(y-2x). \end{aligned}$$

(ii) Using the chain rule we have

$$\begin{aligned} \frac{\partial F}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} \\ &= (\sin u + (u-v)\cos u)\,2x + (-\sin u + e^v)(-2) \\ &= (2x+2)\sin\left(x^2+y\right) + 2x\left(x^2+2x\right)\cos\left(x^2+y\right) - 2\exp(y-2x), \end{aligned}$$

and

$$\frac{\partial F}{\partial y} =$$

■

Theorem 23 ***(Chain Rule)*** *Let $F(t) = f(u(t), v(t))$ with u and v differentiable and f being continuously differentiable in each variable. Then*

$$\frac{dF}{dt} = \frac{\partial f}{\partial u}\frac{du}{dt} + \frac{\partial f}{\partial v}\frac{dv}{dt}$$

Proof. (Not examinable) If we change t to $t + \delta t$, then let δu and δv be the corresponding changes in u and v. Then

$$\delta u = \left(\frac{\mathrm{d}u}{\mathrm{d}t} + \varepsilon_1\right)\delta t, \quad \text{and} \quad \delta v = \left(\frac{\mathrm{d}v}{\mathrm{d}t} + \varepsilon_2\right)\delta t,$$

where $\varepsilon_1, \varepsilon_2 \to 0$ as $\delta t \to 0$. Now

$$\begin{aligned}\delta F &= f(u + \delta u, v + \delta v) - f(u, v) \\ &= [f(u + \delta u, v + \delta v) - f(u, v + \delta v)] + [f(u, v + \delta v) - f(u, v)]\end{aligned}$$

By the Mean-value Theorem (Hilary term Analysis) we have

$$\begin{aligned}f(u + \delta u, v + \delta v) - f(u, v + \delta v) &= \delta u \frac{\partial f}{\partial u}(u + \theta_1 \delta u, v + \delta v), \\ f(u, v + \delta v) - f(u, v) &= \delta v \frac{\partial f}{\partial v}(u, v + \theta_2 \delta v),\end{aligned}$$

for some $\theta_1, \theta_2 \in (0, 1)$. By the continuity of f_u and f_v then we have

$$\begin{aligned}\delta u \frac{\partial f}{\partial u}(u + \theta_1 \delta u, v + \delta v) &= \delta u \left(\frac{\partial f}{\partial u}(u, v) + \eta_1\right) \\ \delta v \frac{\partial f}{\partial v}(u, v + \theta_2 \delta v) &= \delta v \left(\frac{\partial f}{\partial v}(u, v) + \eta_2\right)\end{aligned}$$

where $\eta_1, \eta_2 \to 0$ as $\delta u, \delta v \to 0$.
So, putting this all together

$$\begin{aligned}\frac{\delta F}{\delta t} &= \frac{\delta u}{\delta t}\left(\frac{\partial f}{\partial u}(u, v) + \eta_1\right) + \frac{\delta v}{\delta t}\left(\frac{\partial f}{\partial v}(u, v) + \eta_2\right) \\ &= \left(\frac{\mathrm{d}u}{\mathrm{d}t} + \varepsilon_1\right)\left(\frac{\partial f}{\partial u}(u, v) + \eta_1\right) + \left(\frac{\mathrm{d}v}{\mathrm{d}t} + \varepsilon_2\right)\left(\frac{\partial f}{\partial v}(u, v) + \eta_2\right).\end{aligned}$$

Letting $\delta t \to 0$ we get the required result. ■

Corollary 24 *Let $F(x,y) = f(u(x,y), v(x,y))$ with u, v differentiable in each variable and f being continuously differentiable in each. Then*

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

Example 25 *A particle $P(x,y,z)$ moves in three dimensional space on a helix so that at time t*

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

The temperature T at (x,y,z) equals $xy + yz + zx$. Use (i) direct calculation, (ii) the chain rule, to calculate dT/dt.

Solution. (i)

$$\begin{aligned} T(t) &= x(t)\,y(t) + y(t)\,z(t) + z(t)\,x(t) \\ &= \cos t \sin t + t \sin t + t \cos t \end{aligned}$$

So

$$\frac{dT}{dt} = -\sin^2 t + \cos^2 t + \sin t + \cos t + t\cos t - t\sin t.$$

(ii) Alternatively the chain rule says that

$$\frac{dT}{dt} =$$

■

Example 26 *Let $z = f(xy)$. Show that*

$$x\frac{\partial z}{\partial x} - y\frac{\partial z}{\partial y} = 0. \tag{2.1}$$

Example 27 *Conversely show that any solution of (2.1) is of the form $z = f(xy)$.*

Solution. We first make a change of co-ordinates

$$u = y/x, \qquad v = xy.$$

We are aiming to show that z is a function solely in v, or equivalently that $\partial z/\partial u = 0$. By the chain rule

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}.$$

Solving for x and y in terms of u and v we have that

$$x = \sqrt{\frac{v}{u}} \text{ and } y = \sqrt{uv}.$$

Then

$$\begin{aligned}
\frac{\partial z}{\partial u} &= \frac{-1}{2}\sqrt{\frac{v}{u^3}}\frac{\partial z}{\partial x} + \frac{1}{2}\sqrt{\frac{v}{u}}\frac{\partial z}{\partial y} \\
&= \frac{-1}{2}\frac{x^2}{y}\frac{\partial z}{\partial x} + \frac{1}{2}x\frac{\partial z}{\partial y} \\
&= \frac{-1}{2}\frac{x}{y}\left(x\frac{\partial z}{\partial x} - y\frac{\partial z}{\partial y}\right) = 0
\end{aligned}$$

and $z = f(v) = f(xy)$ as required, where f is an arbitrary differentiable function. ■

Example 28 *A particle P moves around on the unit sphere $r = 1$. Find P's velocity $\mathbf{v}(t) = d\mathbf{r}/dt$ in terms of $\phi,\ \theta,\ \dot{\phi},\ \dot{\theta}$ and verify by direct calculation that $\mathbf{v} \bullet \mathbf{r} = 0$.*

Solution. Recall that

$$x = \sin\theta\cos\phi, \qquad y = \sin\theta\sin\phi, \qquad z = \cos\theta.$$

Now

$$\begin{aligned}
\dot{x} &= \cos\theta\cos\phi\,\dot{\theta} - \sin\theta\sin\phi\,\dot{\phi}, \\
\dot{y} &= \cos\theta\sin\phi\,\dot{\theta} + \sin\theta\cos\phi\,\dot{\phi}, \\
\dot{z} &= -\sin\theta\,\dot{\theta}.
\end{aligned}$$

So

$$\begin{aligned}
\mathbf{v}\cdot\mathbf{r} &= \left(\cos\theta\cos\phi\,\dot{\theta} - \sin\theta\sin\phi\,\dot{\phi}\right)\sin\theta\cos\phi \\
&\quad + \left(\cos\theta\sin\phi\,\dot{\theta} + \sin\theta\cos\phi\,\dot{\phi}\right)\sin\theta\sin\phi \\
&\quad + \left(-\sin\theta\,\dot{\theta}.\right)\cos\theta \\
&= \dot{\theta}\left(\cos\theta\cos^2\phi\sin\theta + \cos\theta\sin^2\phi\sin\theta - \cos\theta\sin\theta\right) \\
&\quad + \dot{\phi}\left(-\sin^2\theta\sin\phi\cos\phi + \sin^2\theta\sin\phi\cos\phi\right) \\
&= 0
\end{aligned}$$

■

This is true for any movement on the sphere. We can prove this much more easily by differentiating the vector identity $\mathbf{r} \bullet \mathbf{r} = 1$ to get $2\mathbf{v} \bullet \mathbf{r} = 0$.

To find the particle's acceleration by means of a chain rule we would need the next theorem.

Theorem 29 *(The Second Order Chain Rule)*

Let $F(x,y) = f(u(x,y), v(x,y))$. *Then*

$$\begin{aligned} F_{xx} &= f_u u_{xx} + f_v v_{xx} + f_{uu}(u_x)^2 + 2f_{uv} v_x u_x + f_{vv}(v_x)^2, \\ F_{xy} &= f_u u_{xy} + f_v v_{xy} + f_{uu} u_x u_y + f_{uv}(v_y u_x + v_x u_y) + f_{vv} v_x v_y, \\ F_{yy} &= f_u u_{yy} + f_v v_{yy} + f_{uu}(u_y)^2 + 2f_{uv} v_y u_y + f_{vv}(v_y)^2. \end{aligned}$$

Proof.

$$\begin{aligned} F_{xx} &= (f_u u_x + f_v v_x)_x \\ &= (f_u)_x u_x + (f_v)_x v_x + f_u u_{xx} + f_v v_{xx} \\ &= (f_{uu} u_x + f_{uv} v_x) u_x + (f_{vu} u_x + f_{vv} v_x) v_x + f_u u_{xx} + f_v v_{xx} \\ &= f_u u_{xx} + f_v v_{xx} + f_{uu}(u_x)^2 + 2f_{uv} v_x u_x + f_{vv}(v_x)^2 \end{aligned}$$

and the other results follow similarly. ■

Example 30 *Recall the definition of parabolic planar co-ordinates*

$$x = \frac{u^2 - v^2}{2}, \qquad y = uv.$$

Show that Laplace's equation

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} = 0$$

transforms into the same equation in parabolic co-ordinates.

Solution. From the second order chain rule we have

■

Example 31 *Find all circularly symmetric solutions to Laplace's equation in the plane.*

Solution. When we write Laplace's equation in terms of planar polar co-ordinates it becomes

$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} = 0.$$

To say that a solution is circularly symmetric means that f is solely a function of r. Hence we have

$$\frac{\mathrm{d}^2 f}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}f}{\mathrm{d}r} = 0.$$

This equation has integrating factor r and so we have

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}f}{\mathrm{d}r}\right) = r\frac{\mathrm{d}^2 f}{\mathrm{d}r^2} + \frac{\mathrm{d}f}{\mathrm{d}r} = 0$$

giving

$$\frac{\mathrm{d}f}{\mathrm{d}r} = \frac{A}{r}$$

and so the general solution is

$$f(r) = A\ln r + B$$

■

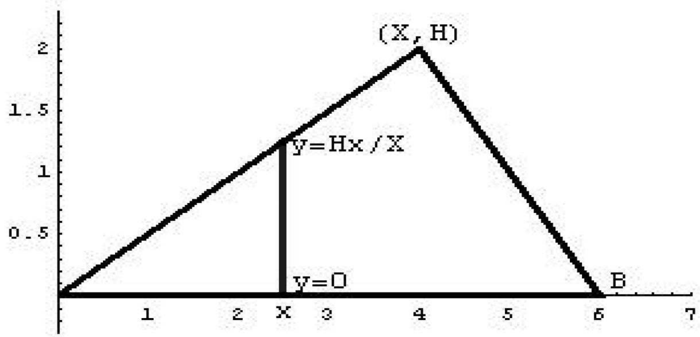
3. JACOBIANS AND AREA

3.1 A Brief Introduction to Double Integrals

In a moment we will meet the definition of a Jacobian. The definition is simple enough, but, in order to motivate the idea behind it, we will need to calculate some areas. As we change from one co-ordinate system to another the equations of curves change, the co-ordinates of points, the rules for functions. Likewise the Jacobian is a rule for how measuring areas, or volumes, changes with a co-ordinate change. Our first two examples of calculating areas involves two examples of *double integrals.*

Example 32 *Calculate the area of the triangle with vertices* $(0,0)\,,(B,0)$ *and* (X,H).

Solution.



We have known, since early school days that the answer is $\frac{1}{2}BH$, but we shall demonstrate this here by means of a double integral. The three bounding lines of the triangles are

$$y = 0, \qquad y = \frac{H}{X}x, \qquad y = \frac{H}{X-B}\left(x - B\right).$$

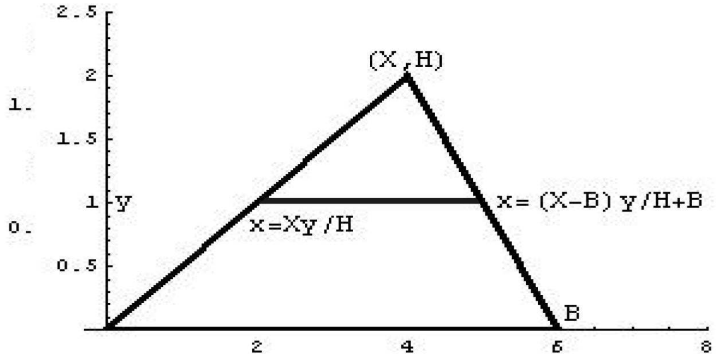
We'll assume here that $0 < X < B$. In order to "pick up" all of the triangle's area we need to let x range from 0 to B and y range appropriately from 0 up to the bounding line above $(x, 0)$. As our x and y vary over the triangle we need to pick up an infinitesimal piece of area $\mathrm{d}x\ \mathrm{d}y$ at each point. We can then calculate the triangle's area as

$$\begin{aligned}
A &= \int_{x=0}^{x=X}\int_{y=0}^{y=Hx/X} \mathrm{d}y\,\mathrm{d}x + \int_{x=X}^{x=B}\int_{y=0}^{y=H(x-B)/(X-B)} \mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=0}^{x=X} \frac{Hx}{X}\,\mathrm{d}x + \int_{x=X}^{x=B} \frac{H\left(x-B\right)}{X-B}\,\mathrm{d}x \\
&= \frac{H}{X}\left[\frac{x^2}{2}\right]_0^X + \frac{H}{X-B}\left[\frac{\left(x-B\right)^2}{2}\right]_X^B \\
&= \frac{H}{X}\frac{X^2}{2} - \frac{H}{X-B}\frac{\left(X-B\right)^2}{2} \\
&= \frac{HX}{2} - \frac{H\left(X-B\right)}{2} \\
&= \frac{HB}{2}.
\end{aligned}$$

■

Integrating this way, y first then x, we would need to treat $X < 0$ and $X > B$ as two further cases. Alternatively, one could also pick up the area by first letting y range from 0 to H and letting x range over the interior points of the triangle at height y.

Solution.



This method is somewhat better as we don't have to treat the three cases of $X < 0$, $0 < X < B$ and $B < X$ separately.

$$\begin{aligned} A &= \int_{y=0}^{y=H} \int_{x=Xy/H}^{x=y(X-B)/H+B} \mathrm{d}x\,\mathrm{d}y \\ &= \int_{y=0}^{y=H} \left(\frac{(X-B)\,y}{H} + B - \frac{yX}{H} \right) \mathrm{d}y \\ &= \int_{y=0}^{y=H} \left(B - \frac{By}{H} \right) \mathrm{d}y \\ &= B \left[y - \frac{y^2}{2H} \right]_{y=0}^{y=H} \\ &= B \left(H - \frac{H^2}{2H} \right) \\ &= \frac{BH}{2} \end{aligned}$$

■

Example 33 *Calculate the area of the disc* $x^2 + y^2 \leqslant a^2$.

Solution. Again we know the answer, namely πa^2.

■

In the first example of the triangle using y as the outside variable and x the inside avoided considering a number of separate cases. For this example of the disc though it would have seemed much more natural to use polar co-ordinates — *if* we knew how to calculate areas with such!

3.2 Jacobians

You may be aware that the modulus of a determinant of a 2×2 matrix A is a measure of how much the map $\mathbf{x} \longmapsto A\mathbf{x}$ stretches area. Note that the extent to which the mapping $\mathbf{x} \longmapsto A\mathbf{x}$ stretches area does not vary from point to point.

The Jacobian, or rather its modulus, is a measure of how a general mapping stretches space locally, near a particular point, even when this stretching effect varies from point to point.

The Jacobian takes its name from the German mathematician Carl Jacobi (1804-1851).

Definition 34 *Given two co-ordinates* $u(x,y)$ *and* $v(x,y)$ *which depend on variables* x *and* y*, we define the* ***Jacobian***

$$\frac{\partial(u,v)}{\partial(x,y)}$$

to be the determinant

$$\det\begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}.$$

Example 35 *Let* u *and* v *be related to* x *and* y *by the matrix map*

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}. \tag{3.1}$$

Then

$$\frac{\partial(u,v)}{\partial(x,y)} = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad-bc.$$

Consider the parallelogram that is the image under the above map of the square with vertices $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$.

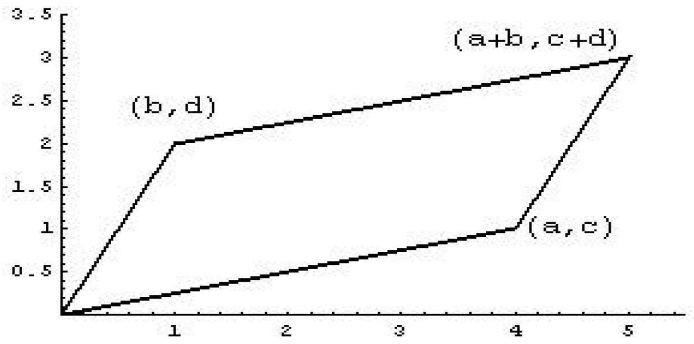


Figure 3-1 Image of the unit square under the map given by relation (3.1)

The area of a parallelogram with sides $\mathbf{a}$ *and* $\mathbf{b}$ *is* $|\mathbf{a}\wedge\mathbf{b}|$*. Hence the above parallelogram's area is*

$$|(a\mathbf{i}+c\mathbf{j})\wedge(b\mathbf{i}+d\mathbf{j})| =$$

This example illustrates that the modulus of the general results that the Jacobian does indeed give the relative change in area under a 2D mapping.

Example 36 *Let $x = r\cos\theta$ and $y = r\sin\theta$ where r and θ are polar co-ordinates. Then*

$$\frac{\partial(x,y)}{\partial(r,\theta)} =$$

Example 37 *In reverse, $r = \sqrt{x^2+y^2}$ and $\theta = \tan^{-1}(y/x)$ and*

$$\begin{aligned}
\frac{\partial(r,\theta)}{\partial(x,y)} &= \det\begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{pmatrix} \\
&= \det\begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix} \\
&= \frac{x^2+y^2}{(x^2+y^2)^{3/2}} \\
&= \frac{1}{\sqrt{x^2+y^2}} = \frac{1}{r}
\end{aligned}$$

The Jacobian satisfies its own sort of chain rule as one might expect:

Proposition 38 *Let r and s be functions of variables u and v which in turn are functions of x and y. Then*

$$\frac{\partial(r,s)}{\partial(x,y)} = \frac{\partial(r,s)}{\partial(u,v)}\frac{\partial(u,v)}{\partial(x,y)}$$

Proof.

$$\begin{aligned}
\frac{\partial(r,s)}{\partial(x,y)} &= \det\begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \end{pmatrix} \\
&= \det\begin{pmatrix} \frac{\partial r}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial r}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial r}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial r}{\partial v}\frac{\partial v}{\partial y} \\ \frac{\partial s}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial s}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial s}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial s}{\partial v}\frac{\partial v}{\partial y} \end{pmatrix} \\
&= \det\left\{\begin{pmatrix} \frac{\partial r}{\partial u} & \frac{\partial r}{\partial v} \\ \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \end{pmatrix}\begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}\right\} \\
&= \frac{\partial(r,s)}{\partial(u,v)}\frac{\partial(u,v)}{\partial(x,y)}.
\end{aligned}$$

■

Corollary 39

$$\frac{\partial(x,y)}{\partial(u,v)}\frac{\partial(u,v)}{\partial(x,y)} = 1$$

Solution. Take $r = x$ and $s = y$ in the previous proposition. Indeed the stronger result holds true that

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}\begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = I_2.$$

■

Example 40 *Recall the definition of parabolic planar co-ordinates*

$$x = \frac{1}{2}\left(u^2 - v^2\right), \qquad y = uv.$$

In this case the Jacobian $\frac{\partial(x,y)}{\partial(u,v)}$ *is given by*

$$\begin{aligned} \frac{\partial(x,y)}{\partial(u,v)} &= \det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \\ &= \det\begin{pmatrix} u & -v \\ v & u \end{pmatrix} \\ &= u^2 + v^2. \end{aligned}$$

Though it is somewhat messy to calculate u and v in terms of x and y we can easily calculate that

$$\begin{aligned} \left(u^2+v^2\right)^2 &= u^4 + 2u^2v^2 + v^4 \\ &= \left(u^2-v^2\right)^2 + 4u^2v^2 \\ &= 4x^2 + 4y^2 \end{aligned}$$

Thus $\quad \frac{\partial(u,v)}{\partial(x,y)} =$

Though 3×3 determinants are off-syllabus, included here are the Jacobians of spherical and cylindrical polar co-ordinates.

Example 41 *(Cylindrical Polar Co-ordinates)*

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z.$$

Then

$$\frac{\partial(x,y,z)}{\partial(r,\theta,z)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial r} \end{vmatrix} =$$

Example 42 *(Spherical Polar Co-ordinates)*

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta.$$

Then

$$\begin{aligned}
\frac{\partial(x,y,z)}{\partial(r,\phi,\theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \theta} \end{vmatrix} \\
&= \begin{vmatrix} \sin\theta\cos\phi & -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ \sin\theta\sin\phi & r\sin\theta\cos\phi & r\cos\theta\sin\phi \\ \cos\theta & 0 & -r\sin\theta \end{vmatrix} \\
&= \cos\theta \begin{vmatrix} -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ r\sin\theta\cos\phi & r\cos\theta\sin\phi \end{vmatrix} - r\sin\theta \begin{vmatrix} \sin\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\sin\theta\cos\phi \end{vmatrix} \\
&= \cos\theta r^2\sin\theta\cos\theta\left(-\sin^2\phi - \cos^2\phi\right) - r^2\sin^3\theta\left(\cos^2\phi + \sin^2\phi\right) \\
&= -r^2\sin\theta\left(\cos^2\theta + \sin^2\theta\right) \\
&= -r^2\sin\theta.
\end{aligned}$$

Definition 43 *Let $R \subseteq \mathbb{R}^2$. Then we define the* ***area*** *of R to be*

$$A(R) = \iint_{(x,y)\in R} dx\, dy.$$

Quite what this definition formally means, and for what regions of R it makes sense to talk of their "area", is actually a very complicated question. But for the simple examples we shall give this definition will be clear and their area unambiguous.

Theorem 44 *Let $f : R \to S$ be a bijection between two regions of $\mathbb{R}^2$, and write $(u, v) = f(x, y)$. Suppose that*

$$\frac{\partial (u, v)}{\partial (x, y)}$$

is defined and non-zero everywhere. Then

$$A(S) = \iint_{(u,v)\in S} du\, dv = \iint_{(x,y)\in R} \left| \frac{\partial (u, v)}{\partial (x, y)} \right| dx\, dy$$

Example 45 *If we return to the case of calculating the area of the disc $x^2 + y^2 \leqslant a^2$, now using polar co-ordinates we have a much simpler double integral. The interior of the disc, in polar co-ordinates, is given by*

$$0 \leqslant r \leqslant a, \qquad 0 \leqslant \theta < 2\pi.$$

So

$$\begin{aligned} A &= \iint_{x^2+y^2\leqslant a^2} dx\, dy = \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=2\pi} \left| \frac{\partial (x, y)}{\partial (r, \theta)} \right| d\theta\, dr \\ &= \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=2\pi} r\, d\theta\, dr \\ &= \int_{r=0}^{r=a} [\theta r]_{\theta=0}^{\theta=2\pi}\, dr \\ &= 2\pi \int_{r=0}^{r=a} r\, dr \\ &= 2\pi \left[\frac{r^2}{2} \right]_{r=0}^{r=a} = \pi a^2. \end{aligned}$$

Proof. (Sketch proof of preceding theorem) Consider the small element of area that is bounded by the co-ordinate lines $u = u_0$ and $u = u_0 + \delta u$ and $v = v_0$ and $v = v_0 + \delta v$. Let's say that $f(x_0, y_0) = (u_0, v_0)$ and consider small changes δx and δy in x and y respectively. We have

$$\begin{aligned}\delta u &= f(x_0 + \delta x, y_0) - f(x_0, y_0) \approx \frac{\partial f}{\partial x}(x_0, y_0)\,\delta x,\\ \delta v &= f(x_0, y_0 + \delta y) - f(x_0, y_0) \approx \frac{\partial f}{\partial y}(x_0, y_0)\,\delta y.\end{aligned}$$

As f takes values in $\mathbb{R}^2$ then the above are vectors in $\mathbb{R}^2$. The area of a parallelogram in $\mathbb{R}^2$ with sides $\mathbf{a}$ and $\mathbf{b}$ is $|\mathbf{a} \wedge \mathbf{b}|$ where $\wedge$ denotes the vector product. So the element of area we are considering is

$$\left|\frac{\partial f}{\partial x}\,\delta x \wedge \frac{\partial f}{\partial y}\,\delta y\right| = \left|\frac{\partial f}{\partial x} \wedge \frac{\partial f}{\partial y}\right|\,\delta x\,\delta y.$$

Now $f = (u, v)$, so $f_x = (u_x, v_x)$, $f_y = (u_y, v_y)$ and

$$\frac{\partial f}{\partial x} \wedge \frac{\partial f}{\partial y} = \left(\frac{\partial u}{\partial x}, \frac{\partial v}{\partial x}\right) \wedge \left(\frac{\partial u}{\partial y}, \frac{\partial v}{\partial y}\right) = \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right)\mathbf{k}.$$

Finally

$$\begin{aligned}\delta A &\approx \left|\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right|\,\delta x\,\delta y.\\ &= \left|\frac{\partial(u, v)}{\partial(x, y)}\right|\,\delta x\,\delta y.\end{aligned}$$

■

Example 46 *A **shear** parallel to the x-axis is a map of the form*

$$(x, y) \longmapsto (x + ky, y).$$

Note that this can also be written as

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and that the determinant in question is 1, so that shears are area-preserving.

If we consider the effect of the shear

$$\begin{pmatrix} 1 & -X/H \\ 0 & 1 \end{pmatrix}$$

on the co-ordinates $(0,0), (X,H), (B,0)$ *of the vertices of the triangle in Example 32 we see that they are moved to* $(0,0), (0,H)$,$(B,0)$ *.So we see that we can calculate the general triangle's area by just considering the special case when* $X = 0$.

*Indeed, if we were so inclined, we could note that the **stretch***

$$\begin{pmatrix} 1/B & 0 \\ 0 & 1/H \end{pmatrix}$$

has determinant $1/(BH)$ *and moves the triangle to* $(0,0), (1,0), (0,1)$. *Hence the original triangle has area* BH *times this simpler triangle.*

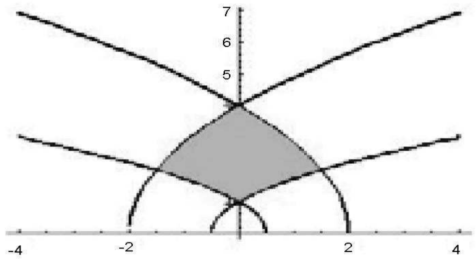
Exercise 47 *By considering an appropriate stretch find the area of the ellipse*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Example 48 *Calculate the area bounded by the curves*

$$2x = 1 - y^2, \qquad 2x = y^2 - 1, \qquad 8x = 16 - y^2, \qquad 8x = y^2 - 16.$$

as shown in the diagram below.



Solution. We see, if we change to planar polar co-ordinates

$$x = \frac{1}{2}\left(u^2 - v^2\right), \qquad y = uv,$$

that the region in question is $1 \leqslant u \leqslant 2, \quad 1 \leqslant v \leqslant 2$. Hence the area is given by

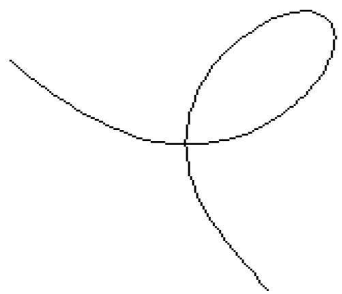
$$\begin{aligned} A &= \int_{u=1}^{u=2} \int_{v=1}^{v=2} \left| \frac{\partial (x,y)}{\partial (u,v)} \right| \mathrm{d}v \, \mathrm{d}u \\ &= \int_{u=1}^{u=2} \int_{v=1}^{v=2} \left(u^2 + v^2\right) \mathrm{d}v \, \mathrm{d}u \\ &= \int_{u=1}^{u=2} \left[u^2 v + \frac{v^3}{3} \right]_{v=1}^{v=2} \mathrm{d}u \\ &= \int_{u=1}^{u=2} \left(u^2 + \frac{7}{3} \right) \mathrm{d}u \\ &= \left[\frac{u^3}{3} + \frac{7u}{3} \right]_1^2 = \frac{7}{3} + \frac{7}{3} = \frac{14}{3}. \end{aligned}$$

■

Example 49 *The area of the sector bounded by the half-lines $\theta = \alpha$ and $\theta = \beta$ and the polar curve $r = R(\theta)$ equals*

$$\begin{aligned} A &= \int_{\theta=\alpha}^{\theta=\beta} \int_{r=0}^{r=R(\theta)} r \ dr\, d\theta \\ &= \int_{\theta=\alpha}^{\theta=\beta} \left[\frac{r^2}{2} \right]_0^{R(\theta)} d\theta \\ &= \frac{1}{2} \int_{\theta=\alpha}^{\theta=\beta} R(\theta)^2 \ d\theta \end{aligned}$$

Example 50 *Find the area bounded by the Folium of Descartes (see diagram) which has equation $x^3 + y^3 = 3xy$.*



Example 51 *Moments of Inertia of a Disc (Off-syllabus)*

The Moment of Inertia I of a set of masses about an axis equals $\sum_i m_i r_i^2$ where m_i is the mass and r_i is the distance of the mass m_i from the axis. Or in the case of a continuous body R, this is given by

$$\int_R r^2 \, dm.$$

So to work out the moment of inertia of a uniform disc of mass M and radius a about a central axis, we need to integrate $r^2\,(\rho r\,dr\,d\theta)$ over the disc — here ρ is the density of the disc so that $M = \rho\pi a^2$. Hence

$$I = \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=2\pi} \rho r^3 \, d\theta \, dr = 2\pi\rho \left[\frac{r^4}{4}\right]_0^a = \frac{\pi\rho a^4}{4} = \frac{Ma^2}{2}.$$

To work out the moment of inertia of the disc about a point on its edge we need to take the origin as this point. The equation of the disc is $(x-a)^2 + y^2 = a^2$ which in polar co-ordinates becomes $r = 2a\cos\theta$.Hence the moment of inertia is

$$\begin{aligned}
I &= \int_{\theta=-\pi/2}^{\theta=\pi/2} \int_{r=0}^{r=2a\cos\theta} \rho r^3 \, dr \, d\theta \\
&= \rho a^4 \int_{\theta=-\pi/2}^{\theta=\pi/2} \frac{(2a\cos\theta)^4}{4} \, d\theta \\
&= \rho a^4 \int_{\theta=-\pi/2}^{\theta=\pi/2} (1+\cos 2\theta)^2 \, d\theta \\
&= \rho a^4 \int_{\theta=-\pi/2}^{\theta=\pi/2} \left\{ \frac{3}{2} + 2\cos 2\theta + \frac{1}{2}\cos 4\theta \right\} d\theta \\
&= \rho a^4 \left[\frac{3\theta}{2} + \sin 2\theta + \frac{1}{8}\sin 4\theta \right]_{\theta=-\pi/2}^{\theta=\pi/2} \\
&= \frac{3\pi\rho a^4}{2} = \frac{3Ma^2}{2}
\end{aligned}$$

Example 52 *To calculate the volume of the sphere $x^2 + y^2 + z^2 \leqslant a^2$, it seems most natural to use spherical polar co-ordinates. The interior of the sphere, in spherical polar co-ordinates, is given by*

$$0 \leqslant r \leqslant a, \qquad 0 \leqslant \phi < 2\pi, \qquad 0 \leqslant \theta < \pi.$$

So

$$A = \iiint_{x^2+y^2+z^2 \leqslant a^2} dx\, dy\, dz =$$

Exercise 53 *Find the volume of the ellipsoid*

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

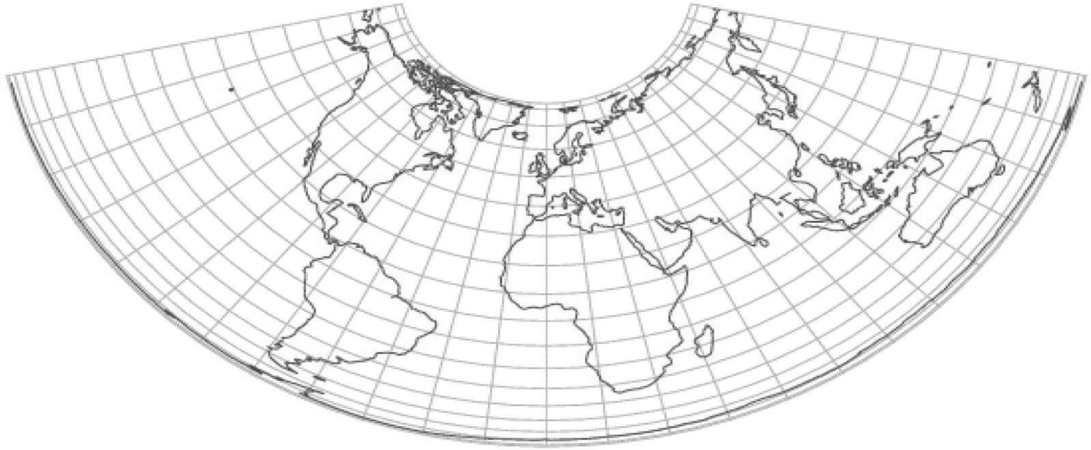
Example 54 *Equal Area Projections*
Recall the definition of spherical polar co-ordinates

$$x = r \sin\theta \cos\phi, \qquad y = r \sin\theta \sin\phi, \qquad z = r\cos\theta,$$

whose Jacobian has magnitude

$$\left| \frac{\partial (x, y, z)}{\partial (r, \phi, \theta)} \right| = r^2 \sin\theta$$

where $\frac{\pi}{2} - \theta$ *is latitude and* ϕ *is longitude. If then we are considering the element of area on the planet's surface* $r = a$*, represented by* θ *and* ϕ*, then the local area distortion is* $a^2 \sin\theta$*.*



The Albers Equal-area Conic Projection is an equal-area projection. In terms of θ *and* ϕ *the projection is*

$$x = \frac{1}{n}\sqrt{C - 2n\cos\theta}\sin n\phi, \qquad y = \rho_0 - \frac{1}{n}\sqrt{C - 2n\cos\theta}\cos n\phi,$$

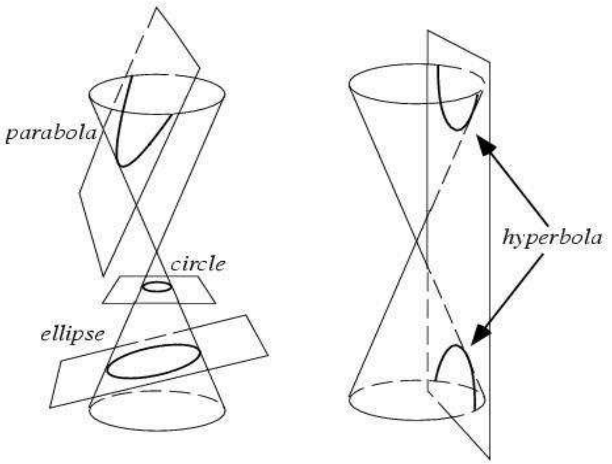
where n, C, ρ_0 *are constants. Then the Jacobian of the projection is*

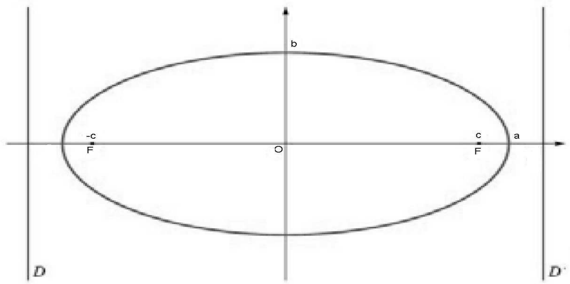
$$\begin{aligned} \left|\frac{\partial(x,y)}{\partial(\phi,\theta)}\right| &= \left|\det\begin{pmatrix} \sqrt{C - 2n\cos\theta}\cos n\phi & \frac{\sin\theta}{\sqrt{C-2n\cos\theta}}\sin n\phi \\ \sqrt{C - 2n\cos\theta}\sin n\phi & -\frac{2n\sin\theta}{\sqrt{C-2n\cos\theta}}\cos n\phi \end{pmatrix}\right| \\ &= \cos^2 n\phi\sin\theta + \sin^2 n\phi\sin\theta \\ &= \sin\theta. \end{aligned}$$

4. THE GRADIENT VECTOR

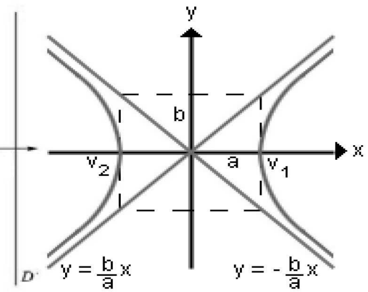
4.1 Standard Curves and Surfaces

In the next section we will meet the *gradient vector* ∇f of a scalar function f. The gradient vector is a way of calculating the normal of a curve, or surface, $f = c$ defined by a function. Here we introduce a few standard examples of some curves and surfaces





ellipse



hyperbola

Example 55 *(Conic Sections) Conic sections, also referred to as conics, are non-degenerate curves formed by the intersection of a plane with a cone. Equivalently, they correspond to curves in a 2D plane which are solutions of algebraic equations of degree 2. The standard form of equation of the conics (circle, ellipse, parabola, hyperbola) are*

- ***circle:*** $x^2 + y^2 = a^2$; *parameterisation* $(a\cos\theta, a\sin\theta)$; *area* πa^2.
- ***ellipse:*** $x^2/a^2 + y^2/b^2 = 1$ $(b < a)$; *parameterisation* $(a\cos\theta, b\sin\theta)$; *eccentricity* $e^2 = 1 - b^2/a^2$; *foci* $(\pm ae, 0)$; *directrices* $x = \pm a/e$; *area* πab.
- ***parabola:*** $y^2 = 4ax$; *parameterisation* $\left(at^2, 2at\right)$; *eccentricity* $e = 1$; *focus* $(a, 0)$; *directrix* $x = -a$.
- ***hyperbola:*** $x^2/a^2 - y^2/b^2 = 1$; *parameterisation* $(a\sec t, b\tan t)$ *or* $(\pm a\cosh t, b\sinh t)$; *eccentricity* $e^2 = 1 + b^2/a^2$; *foci* $(\pm ae, 0)$; *directrices* $x = \pm a/e$; *asymptotes* $y = \pm bx/a$

Example 56 *Find the point of the ellipse* $x^2/4 + y^2/9 = 1$ *which is closest to the point* $(3, 3)$.

In order to solve this we can turn the question into a one-variable maximization problem by parameterising the ellipse as $(2\cos t, 3\sin t)$. *The distance* $D(t)$ *of this point from* $(3, 3)$ *is given by*

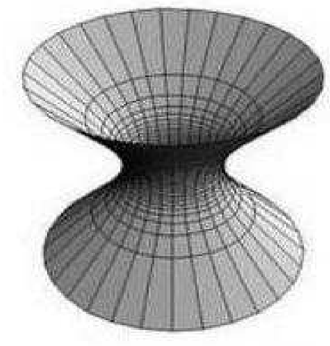
$$D(t)^2 = (3 - 2\cos t)^2 + (3 - 3\sin t)^2 .$$

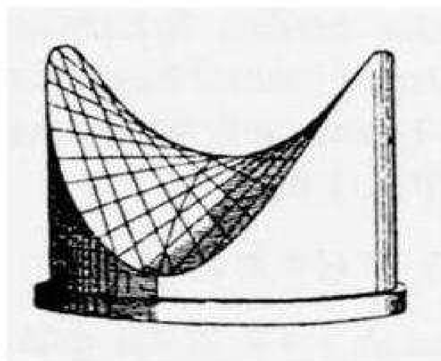
Thus ...

Exercise 57 *To find the point on the ellipse closest to a* ***line*** *we would need to find the point where the tangent to the ellipse is parallel to the line.*

Example 58 *(Examples of Quadrat*
The standard form of equation of th

- ***Sphere:*** $x^2 + y^2 + z^2 = a^2$;
- ***Ellipsoid:*** $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$;
- ***Hyperboloid of One Sheet:*** $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$;
- ***Hyperboloid of Two Sheets:*** $x^2/a^2 - y^2/b^2 - z^2/c^2 = 1$;
- ***Paraboloid:*** $z = x^2 + y^2$;
- ***Hyperbolic Paraboloid:*** $z = x^2 - y^2$;
- ***Cone:*** $x^2 + y^2 = z^2$.


one sheet hyperboloid


hyperbolic paraboloid

Problem 59 *How would one determine the closest point on a surface to: another point? a line? a plane?*

Example 60 *Find the tangent and normal to the ellipse $x^2/a^2 + y^2/b^2 = 1$ at the point (X, Y).*

Solution. One way would be to implicitly differentiate the equation and obtain

$$\frac{2x}{a^2} + \frac{2y}{b^2}\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

giving $\mathrm{d}y/\mathrm{d}x = -b^2X/\left(a^2Y\right)$ at the point in question. Then the normal has gradient $a^2Y/\left(b^2X\right)$. And so the two equations are

$$\begin{aligned} y - Y &= \frac{-b^2X}{a^2Y}(x - X) \qquad \text{(tangent)} \\ y - Y &= \frac{a^2Y}{b^2X}(x - X) \qquad \text{(normal)} \end{aligned}$$

The tangent's equation simplifies somewhat to

$$b^2Xx + a^2Yy = a^2b^2.$$

Alternatively we know that a parameterisation of the ellipse is given by

$$\mathbf{r}(t) = (a\cos t, b\sin t).$$

So a **tangent vector** to the ellipse at $\mathbf{r}(t)$ equals

This is a vector in the tangent direction. Hence a vector in the normal direction is

■

We probably all feel we know a smooth surface in $\mathbb{R}^3$ when we see one, and this instinct for what a surface is will be satisfactory for the purposes of this course. We recognise all the previous examples as smooth surfaces, with the exception of the cone at $(0,0,0)$ and could well be comfortable working out their tangent planes and normals. In the main we will be happy to work with surfaces without a formal definition, but for those seeking a more rigorous treatment of the topic, the following might provide a suitable working definition.

Definition 61 *(Off-syllabus) A* ***smooth parameterised surface*** *is a map* $\mathbf{r}$*, known as the* ***parameterisation***

$$\mathbf{r} : U \to \mathbb{R}^3 : (u,v) \mapsto (x(u,v), y(u,v), z(u,v))$$

from an ***open subset*** $U \subseteq \mathbb{R}^2$ *to* $\mathbb{R}^3$ *such that*

- x, y, z *have continuous partial derivatives with respect to* u *and* v *of all orders*
- $\mathbf{r}$ *is a bijection with both* $\mathbf{r}$ *and* $\mathbf{r}^{-1}$ *being continuous*
- *at each point the vectors*

$$\frac{\partial \mathbf{r}}{\partial u} \quad \text{and} \quad \frac{\partial \mathbf{r}}{\partial v}$$

  are linearly independent (i.e. are not scalar multiples of one another).

We will not be looking to treat this definition with any generality. Rather we shall just look to parameterise some of the "standard" surfaces previously described and calculate some tangents and normals. We define:

Definition 62 *Let* $\mathbf{r} : U \to \mathbb{R}^3$ *be a smooth parameterised surface and let* $\mathbf{p}$ *be a point on the surface. The plane containing* $\mathbf{p}$ *and which is parallel to the vectors*

$$\frac{\partial \mathbf{r}}{\partial u}(\mathbf{p}) \quad \text{and} \quad \frac{\partial \mathbf{r}}{\partial v}(\mathbf{p})$$

is called the ***tangent plane*** *to* $\mathbf{r}(U)$ *at* $\mathbf{p}$*. Because these vectors are independent the tangent plane is well-defined.*

Definition 63 *Any vector in the direction*

$$\frac{\partial \mathbf{r}}{\partial u}(\mathbf{p}) \wedge \frac{\partial \mathbf{r}}{\partial v}(\mathbf{p})$$

is said to be ***normal*** *to the surface at* $\mathbf{p}$*. There are two* ***unit normals*** *of length one.*

Example 64 *We have already met a parameterisation for the sphere* $x^2 + y^2 + z^2 = a^2$ *with spherical polar co-ordinates. This given by*

$$\mathbf{r}(\phi, \theta) = (a \sin\theta \cos\phi, a \sin\theta \sin\phi, a \cos\theta).$$

We already know the outward-pointing unit normal at $\mathbf{r}(\theta, \phi)$ *is* $\mathbf{r}(\theta, \phi)/a$ *but let's verify this with the previous definitions and find the tangent plane. We have*

$$\frac{\partial \mathbf{r}}{\partial \phi} =$$

$$\frac{\partial \mathbf{r}}{\partial \theta} =$$

Hence

$$\frac{\partial \mathbf{r}}{\partial \phi} \wedge \frac{\partial \mathbf{r}}{\partial \theta} =$$

and so the two unit normals are $\pm(\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta)$*. The tangent plane at* $\mathbf{r}(\phi, \theta)$ *is then*

$$\mathbf{r} \bullet (\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta) = a.$$

Example 65 *Find the normal and tangent plane to the point (X, Y, Z) on the hyperbolic paraboloid $z = x^2 - y^2$.*

Solution. This surface has a simple choice of parameterisation as there is exactly one point lying above, or below, the point $(x, y, 0)$. So we can take a parameterisation

$$\mathbf{r}(x, y) = \left(x, y, x^2 - y^2\right).$$

We then have

$$\frac{\partial \mathbf{r}}{\partial x} = (1, 0, 2x), \qquad \frac{\partial \mathbf{r}}{\partial y} = (0, 1, -2y).$$

So

$$\frac{\partial \mathbf{r}}{\partial x} \wedge \frac{\partial \mathbf{r}}{\partial y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 2x \\ 0 & 1 & -2y \end{vmatrix} = \begin{pmatrix} -2x \\ 2y \\ 1 \end{pmatrix}.$$

A normal vector to the surface at (X, Y, Z) is then $(-2X, 2Y, 1)$ and we see that the equation of the tangent plane is

$$\begin{aligned} \mathbf{r} \bullet (-2X, 2Y, 1) &= \left(X, Y, X^2 - Y^2\right) \bullet (-2X, 2Y, 1) \\ &= -2X^2 + 2Y^2 + X^2 - Y^2 \\ &= Y^2 - X^2 = -Z \end{aligned}$$

or equivalently

$$2Xx - 2Yy + z = Z.$$

■

4.2 The Gradient Vector

Definition 66 *Given a scalar function $f : \mathbb{R}^n \to \mathbb{R}$, whose partial derivatives all exist, the* ***gradient vector***

$$\nabla f \qquad \text{or} \qquad \text{grad } f$$

is defined as

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right).$$

The symbol ∇ is pronounced "grad" usually, but also "del" or "nabla".

Example 67 *Let*

$$f(x, y, z) = 2xy^2 + ze^x + yz.$$

Then

$$\nabla f = \left(2y^2 + ze^x, 4xy + z, e^x + y\right).$$

Example 68 *Let*

$$g(x, y, z) = x^2 + y^2 + z^2.$$

Then

$$\nabla g = (2x, 2y, 2z).$$

Note that ∇g is normal to the spheres $g =$ const.

Example 69 *Consider the vector field*

$$\mathbf{v}(x, y, z) = \left(2xy + z\cos(zx), x^2 + e^{y-z}, x\cos(zx) - e^{y-z}\right).$$

Find all the scalar functions $f(x, y, z)$ such that $\mathbf{v} = \nabla f$.

Solution.

∎

Exercise 70 *Show that there is no function* $f(x, y, z)$ *such that* $\nabla f(x, y, z) = (y, z, x)$.

Proposition 71 *(Planar Polar Co-ordinates) Let*

$$\mathbf{e}_r = (\cos\theta, \sin\theta), \qquad \mathbf{e}_\theta = (-\sin\theta, \cos\theta).$$

These are perpendicular unit vectors, pointing along the curves $r =$ const. and $\theta =$ const. Then

$$\nabla\psi = \frac{\partial\psi}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial\psi}{\partial\theta}\mathbf{e}_\theta.$$

Proof. It is easy to check that

$$\begin{aligned} \mathbf{e}_r \bullet \mathbf{e}_r &= \cos^2\theta + \sin^2\theta = 1, \\ \mathbf{e}_\theta \bullet \mathbf{e}_\theta &= \sin^2\theta + \cos^2\theta = 1, \\ \mathbf{e}_r \bullet \mathbf{e}_\theta &= -\sin\theta\cos\theta + \sin\theta\cos\theta = 0. \end{aligned}$$

The line $\theta = c$ is parameterised by $\theta(r) = (r\cos c, r\sin c)$ and we have $\theta'(r) = \mathbf{e}_r$; similarly $r(\theta) = (a\cos\theta, a\sin\theta)$ and we see $r'(\theta) = a\mathbf{e}_\theta$. Recall that the relationship between r, θ and x, y are given by

$$\begin{aligned} r &= \sqrt{x^2+y^2}, \qquad \frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2}} = \cos\theta, \qquad \frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2+y^2}} = \sin\theta, \\ \theta &= \tan^{-1}\frac{y}{x}, \qquad \frac{\partial\theta}{\partial x} = \frac{-y}{x^2+y^2} = \frac{-\sin\theta}{r}, \qquad \frac{\partial\theta}{\partial y} = \frac{x}{x^2+y^2} = \frac{\cos\theta}{r}, \end{aligned}$$

and so

$$\begin{aligned} \nabla\psi &= \frac{\partial\psi}{\partial x}\mathbf{i} + \frac{\partial\psi}{\partial y}\mathbf{j} \\ &= \end{aligned}$$

■

Exercise 72 *(Spherical Polar Co-ordinates)*
(i) Show that

$$\begin{aligned}\mathbf{e}_r &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,.\\ \mathbf{e}_\phi &= (-\sin\phi, \cos\phi)\\ \mathbf{e}_\theta &= (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta)\end{aligned}$$

are mutually perpendicular unit vectors.
(ii) Show that if $\mathbf{r}(t)$ *is a vector with co-ordinates* $r(t), \phi(t), \theta(t)$ *at time* t *then*

$$\frac{d\mathbf{r}}{dt} = \dot{r}\mathbf{e}_r + \dot{\phi}\left(r\sin\theta\mathbf{e}_\phi\right) + \dot{\theta}\left(re_\theta\right).$$

(iii) Show further that

$$\nabla f = \frac{\partial f}{\partial r}\mathbf{e}_r + \frac{1}{r\sin\theta}\frac{\partial f}{\partial\phi}\mathbf{e}_\phi + \frac{1}{r}\frac{\partial f}{\partial\theta}\mathbf{e}_\theta.$$

Example 73 *The* ***gravitational potential*** *at a distance r from a point mass M is given by*

$$V = -\frac{GM}{r}.$$

Using spherical co-ordinates we can write down straight away that

$$\nabla V = \frac{GM}{r^2}\mathbf{e}_r.$$

Alternatively using Cartesian co-ordinates we would have $V = GM\left(x^2+y^2+z^2\right)^{-1/2}$ and find

$$\nabla V = \frac{GM}{\left(x^2+y^2+z^2\right)^{3/2}}(x,y,z) = \frac{GM\mathbf{r}}{r^3} = \frac{GM}{r^2}\mathbf{e}_r.$$

By definition the ***gravitational field*** *is the quantity $-\nabla V$.*

Definition 74 *Let $f:\mathbb{R}^n \to \mathbb{R}$ be a differentiable scalar function and let $\mathbf{u}$ be a unit vector. Then the* ***directional derivative*** *of f at $\mathbf{a}$ in the direction $\mathbf{u}$ equals*

$$\lim_{t\to 0}\left(\frac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t}\right).$$

This is the rate of change of the function f at $\mathbf{a}$ in the direction $\mathbf{u}$.

Example 75 *Let $f(x,y,z) = x^2y - z^2$ and let $\mathbf{a} = (1,1,1)$. Then calculate the directional derivative of f at $\mathbf{a}$ in the direction $\mathbf{u} = (u_1,u_2,u_3)$. In what direction does f increase most rapidly?*

Solution. Now

$$\begin{aligned}\frac{f(\mathbf{a}+t\mathbf{u})-f(\mathbf{a})}{t} &= \frac{\left[(1+tu_1)^2(1+tu_2)-(1+tu_3)^2\right]-\left(1^2 1-1^2\right)}{t} \\ &= (2u_1+u_2-2u_3)+\left(u_1^2+2u_1u_2\right)t+u_1^2u_2t^2 \\ &\to 2u_1+u_2-2u_3 \quad \text{as } t\to 0.\end{aligned}$$

This is the directional derivative asked for. What is the largest this can be as we vary $\mathbf{u}$ over all positive unit vectors? Well in the direction $\left(\frac{2}{3},\frac{1}{3},\frac{-2}{3}\right)$ the directional derivative is

$$\frac{4}{3}+\frac{1}{3}+\frac{4}{3}=3.$$

On the other hand

$$2u_1+u_2-2u_3=(2,1,-2)\bullet\mathbf{u}=3\left|\mathbf{u}\right|\cos\theta=3\cos\theta$$

takes a maximum of 3 when $\theta=0$ and $\mathbf{u}=(2/3,1/3,-2/3)$. ■

Proposition 76 *The directional derivative of a function f at the point* $\mathbf{a}$ *in the direction* $\mathbf{u}$ *equals* $\nabla f(\mathbf{a}) \bullet \mathbf{u}$

Proof. Let

$$F(t) = f(\mathbf{a}+t\mathbf{u}) = f(a_1 + tu_1, \ldots, a_n + tu_n).$$

Then

$$\lim_{t\to 0}\left(\frac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t}\right) = \lim_{t\to 0}\left(\frac{F(t) - F(0)}{t}\right) = F'(0).$$

Now

$$\begin{aligned} F'(0) &= \left.\frac{\mathrm{d}F}{\mathrm{d}t}\right|_{t=0} \\ &= \frac{\partial f}{\partial x_1}(\mathbf{a})\frac{\mathrm{d}x_1}{\mathrm{d}t} + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{a})\frac{\mathrm{d}x_n}{\mathrm{d}t} \\ &= \frac{\partial f}{\partial x_1}(\mathbf{a})\, u_1 + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{a})\, u_n \\ &= \nabla f(\mathbf{a}) \bullet \mathbf{u}. \end{aligned}$$

■

Corollary 77 *The rate of change of f is greatest in the direction ∇f, that is when* $\mathbf{u} = \nabla f / |\nabla f|$.

Definition 78 *A* ***level set*** *of a function* $f:\mathbb{R}^3\to\mathbb{R}$ *is a set of points*

$$\left\{(x,y,z)\in\mathbb{R}^3 : f\left(x,y,z\right)=c\right\}$$

where c *is a constant. For suitably "nice" functions* f *and constants* c *the level set is a surface in* $\mathbb{R}^3$. *Note that all the quadrics given in Example 58 are level sets.*

Proposition 79 *Given a surface* $S\subseteq\mathbb{R}^3$ *with equation* $f\left(x,y,z\right)=c$ *and a point* $\mathbf{p}\in S$ *then* $\nabla f\left(\mathbf{p}\right)$ *is normal to* S *at* $\mathbf{p}$.

Proof. Let u and v be co-ordinates near $\mathbf{p}$ and $\mathbf{r}:(u,v)\to\mathbf{r}\left(u,v\right)$ be a parameterisation of part of S. Recall that the normal to S at $\mathbf{p}$ is in the direction

$$\frac{\partial\mathbf{r}}{\partial u}\wedge\frac{\partial\mathbf{r}}{\partial v}.$$

If we write $\mathbf{r}\left(u,v\right)=\left(x\left(u,v\right),y\left(u,v\right),z\left(u,v\right)\right)$ then we see that

$$\begin{aligned}
\nabla f\bullet\frac{\partial\mathbf{r}}{\partial u} &= \left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right)\bullet\left(\frac{\partial x}{\partial u},\frac{\partial y}{\partial u},\frac{\partial z}{\partial u}\right)\\
&= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial u}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial u}\\
&= \frac{\partial f}{\partial u}\qquad\text{[by the chain rule]}\\
&= 0
\end{aligned}$$

as f is constant as a function of u and v. Similarly $\nabla f\bullet\partial\mathbf{r}/\partial v=0$ and hence ∇f is in the direction of $\partial\mathbf{r}/\partial u\wedge\partial\mathbf{r}/\partial v$ and so normal to the surface. ■

Example 80 *In Example 65 we determined the normal at* (X, Y, Z) *to the hyperbolic paraboloid* $z = x^2 - y^2$. *If we introduce the function*

$$f(x.y, z) = x^2 - y^2 - z$$

then we can know the normal is

$$\nabla f(X, Y, Z) = (2X, -2Y, -1)$$

which is find the normal to vector we found previously.

Example 81 *Find the point on the ellipsoid*

$$\frac{x^2}{4} + \frac{y^2}{9} + z^2 = 1$$

which is closest to the plane $x + 2y + z = 10$.

Solution. The normal to the plane is parallel to $(1, 2, 1)$ everywhere. The point on the ellipsoid which is closest will also have normal $(1, 2, 1)$. If we set

$$f(x.y, z) = \frac{x^2}{4} + \frac{y^2}{9} + z^2 - 1,$$

then the gradient equals

$$\nabla f = \left(\frac{x}{2}, \frac{2y}{9}, 2z\right).$$

■

Example 82 *The temperature T in $\mathbb{R}^3$ is given by*

$$T(x, y, z) = x + y^2 - z^2.$$

Given that heat flows in the direction of $-\nabla T$, describe the curve along which heat moves from the point $(1, 1, 1)$.

Solution. We have

$$\nabla T = (1, 2y, -2z),$$

which is the direction in which heat flows. If we parameterise the flow of the heat (say by arc-length s) as $\mathbf{r}(s) = (x(s), y(s), z(s))$ then we have

$$\begin{aligned} \frac{y'(s)}{y(s)} &= 2x'(s), \\ \frac{z'(s)}{z(s)} &= -2x'(s), \end{aligned}$$

and integrating the last two equations we get

$$\begin{aligned} \ln y(s) &= 2x(s) - 2 \\ \ln z(s) &= -2x(s) + 2. \end{aligned}$$

Hence the equation of the heat's path is

$$2x = \ln y + 2 = 2 - \ln z,$$

moving along the direction of $(1, 2, -2)$ from $(1, 1, 1)$. ■

Proposition 83 *Let f and g be functions of x, y, z. Then*

$$\begin{array}{lrl} \textit{(i)} & \nabla(fg) = & f\nabla g + g\nabla f \\ \textit{(ii)} & \nabla(f^n) = & nf^{n-1}\nabla f \\ \textit{(iii)} & \nabla\left(\frac{f}{g}\right) = & \dfrac{g\nabla f - f\nabla g}{g^2} \\ \textit{(iv)} & \nabla(f \circ g)(\mathbf{x}) = & f'(g(\mathbf{x}))\nabla g(\mathbf{x}) \end{array}$$

Proof. See Exercise Sheets for (i), (ii), (iii). The proof of (iv) follows from the chain rule ...

$$\begin{aligned} \nabla(f \circ g)(\mathbf{x}) &= \left(\frac{\partial}{\partial x} f(g(\mathbf{x})), \frac{\partial}{\partial y} f(g(\mathbf{x})), \frac{\partial}{\partial z} f(g(\mathbf{x}))\right) \\ &= \left(f'(g(\mathbf{x}))\frac{\partial g}{\partial x}(\mathbf{x}), f'(g(\mathbf{x}))\frac{\partial g}{\partial y0}(\mathbf{x}), f'(g(\mathbf{x}))\frac{\partial g}{\partial z}(\mathbf{x})\right) \\ &= f'(g(\mathbf{x}))\nabla g(\mathbf{x}) \end{aligned}$$

■

Example 84 *There are two important operations related to grad* ∇*. For a vector function* $\mathbf{v} = (v_1, v_2, v_3)$ *we define*

$$\nabla \wedge \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

This is also called curl $\mathbf{v}$*.*

Secondly

$$\nabla \bullet \mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}$$

which is also called div $\mathbf{v}$*.*

Example 85 *These operations satisfy the identities*

$$\begin{aligned} \nabla \bullet (\nabla \wedge \mathbf{v}) &= 0 \\ \nabla \wedge \nabla f &= \mathbf{0} \end{aligned}$$

for any vector function $\mathbf{v}$ *and scalar function* f*. To prove the latter we see*

$$\begin{aligned} \nabla \wedge \nabla f &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} \\ &= (f_{zy} - f_{yz})\,\mathbf{i} + (f_{xz} - f_{zx})\,\mathbf{j} + (f_{yx} - f_{xy})\,\mathbf{k} = \mathbf{0}. \end{aligned}$$

The Laplacian ∇^2 *given by*

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

can also be rewritten as

$$\nabla^2 f = \text{div (grad } f).$$

Example 86 *Recall in Example 69 we introduced*

$$\mathbf{v}(x,y,z) = \left(2xy + z\cos(zx), x^2 + e^{y-z}, x\cos(zx) - e^{y-z}\right)$$

and noted that $\mathbf{v} = \nabla f$ *where*

$$f(x,y,z) = x^2 y + \sin(zx) + e^{y-z}.$$

From the previous example it follows that $\nabla \wedge \mathbf{v} = \mathbf{0}$.

Another previous exercise was to show that

$$\mathbf{w} = (y, z, x)$$

is not the gradient of any function (i.e. $\nexists \psi$ *for which* $\nabla\psi = \mathbf{w}$*). Note that*

$$\nabla \wedge \mathbf{w} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & z & x \end{vmatrix} = (-1,-1,-1) \neq \mathbf{0}.$$

In fact, the converse of $\nabla \wedge \nabla f = \mathbf{0}$ *does hold. That is, if* $\mathbf{F}$ *is a vector-valued function on* $\mathbb{R}^3$ *such that* $\nabla \wedge \mathbf{F} = \mathbf{0}$ *then there exists a scalar function* ϕ*, known as a* ***potential*** *such that* $\mathbf{F} = \nabla\phi$.

The potential ϕ *can be defined as*

$$\phi(\mathbf{p}) = \int_0^1 \mathbf{F}(t\mathbf{p}) \bullet \mathbf{p}\ dt.$$

[That this is well-defined and satisfies $\mathbf{F} = \nabla\phi$ *follows from Stokes' Theorem — see later courses.] So in the case of* $\mathbf{v}$ *this equals*

$$\begin{aligned} & \int_0^1 \left\{ \left[2t^2xy + tz\cos\left(t^2zx\right)\right] x + \left[t^2x^2 + e^{t(y-z)}\right] y + \left[tx\cos\left(t^2zx\right) - e^{t(y-z)}\right] z \right\} dt \\ = & \left[\frac{2t^3}{3}x^2y + \frac{1}{2}\sin\left(t^2zx\right) + \frac{t^3}{3}x^2y + \frac{ye^{t(y-z)}}{y-z} + \frac{1}{2}\sin\left(t^2zx\right) - \frac{ze^{t(y-z)}}{y-z} \right]_0^1 \\ = & \ x^2y + \sin(zx) + e^{y-z}. \end{aligned}$$

5. CRITICAL POINTS (NON-DEGENERATE)

In this brief chapter we will define, and investigate the properties of, the non-degenerate critical points of a function of two variables. Recall:

Definition 87 *The critical points of a real function of a single variable, $f(x)$, are all values of x for which*

$$f'(x) = 0.$$

Remarks

- The phrase "stationary point" is synonymous with "critical point".
- Critical points of a function of a single variable can be classified as a local maximum, a local minimum or a point of inflection.
- Suppose $f(x)$ has a stationary point at x^0, i.e. $f'(x_0) = 0$. Suppose further that $f''(x_0) < 0$. Then, the stationary point at $x = x_0$ is a local maximum.
- Suppose $f(x)$ has a stationary point at x^0, i.e. $f'(x_0) = 0$. Suppose further that $f''(x_0) > 0$. Then, the stationary point at $x = x_0$ is a local minimum.
- Suppose $f(x)$ has a stationary point at x^0, i.e. $f'(x_0) = 0$. Suppose further that $f''(x_0) = 0$. Then, the stationary point at $x = x_0$ could be a local maximum, a local minimum, or a point of inflection. For example, sketch the graphs and consider the critical points of the functions $f(x) = x^4$, $g(x) = -x^4$, $h(x) = x^3$.

Our aim is to consider the generalisation of the above to functions of two variables.

Definition 88 *The critical points of $f(x, y)$, are all values of (x, y) for which*

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0, \qquad i.e. \qquad \nabla f = \mathbf{0}.$$

Remarks

- Once more, the phrase "stationary point" is synonymous with "critical point".
- Analogously, the critical points of a function of n variables, $f(x_1, ..., x_n)$, are all values of $(x_1, ..., x_n)$ for which $\nabla f = \mathbf{0}$.

Theorem 89 *Suppose $f(x,y)$ has a stationary point at $(x,y)=(a,b)$. Set $f_{xx}(a,b)=r,\ f_{xy}(a,b)=s,\ f_{yy}(a,b)=t$. Then*

(i) If $r<0,\ t<0,\ rt-s^2>0$ then $f(x,y)$ has a local maximum at $(x,y)=(a,b)$.
(ii) If $r>0,\ t>0,\ rt-s^2>0$ then $f(x,y)$ has a local minimum at $(x,y)=(a,b)$.
(iii) If $rt-s^2<0$ then $f(x,y)$ has a saddle point at $(x,y)=(a,b)$.

Remarks

- A saddle point entails a plot of the function $f(x,y)$ in the region immediately around the critical point (a,b) is analogous to a horse's saddle ... the function increases in one direction and decreases in another as one moves away from the stationary point.
- One often refers to $rt-s^2$ as the discriminant.
- Not all values of $r,\ t,\ rt-s^2$ have been considered. In particular, for the values of the variables which have not been considered above there is degeneracy and further information is required to classify the critical point.
  - This is analogous to when $f''(x_0)=0$ at a stationary point for a function of one variable.
  - Degenerate cases will be classified in next term's lectures.

Proof. Let $\mathbf{n}(\theta)=(\cos\theta,\ \sin\theta)$ be the unit vector in the direction θ. We have the directional derivatives

$$\frac{df}{dn}\stackrel{def}{=}\mathbf{n}\cdot\nabla f=\cos\theta f_x+\sin\theta f_y,$$

$$\frac{d^2f}{dn^2}\stackrel{def}{=}\mathbf{n}\cdot\nabla(\mathbf{n}\cdot\nabla f)=\mathbf{n}\cdot\nabla(\cos\theta f_x+\sin\theta f_y)=\cos\theta\frac{\partial}{\partial x}(\cos\theta f_x+\sin\theta f_y)+\sin\theta\frac{\partial}{\partial y}(\cos\theta f_x+\sin\theta f_y)=\cos^2\theta f_{xx}+2\cos\theta\sin\theta f_{xy}+\sin^2\theta f_{yy}.$$

Thus, for θ fixed, $df/dn,\ d^2f/dn^2$ give the first and second derivative for f in the direction $\mathbf{n}(\theta) = (\cos\theta,\ \sin\theta)$. Clearly, df/dn is zero at a critical point. Additionally, at the critical point we have, by definition, $f_{xx}(a,b) = r,\ f_{xy}(a,b) = s,\ f_{yy}(a,b) = t$ and thus

$$\frac{d^2f}{dn^2}(a,b) = r\cos^2\theta + 2s\cos\theta\sin\theta + t\sin^2\theta = r\left(\cos\theta + \frac{s}{r}\sin\theta\right)^2 + \frac{1}{r}\left(rt - s^2\right)\sin^2\theta.$$

(i) Suppose $r < 0,\ t < 0,\ rt - s^2 > 0$. Then, we have that

$$\frac{d^2f}{dn^2}(a,b) < 0 \qquad \forall\theta.$$

Thus, irrespective of the direction θ one moves away from the critical point, the function f *always* decreases. Thus there is a local maximum at the critical point.

(ii) Suppose $r > 0,\ t > 0,\ rt - s^2 > 0$. Then, we have that

$$\frac{d^2f}{dn^2}(a,b) > 0 \qquad \forall\theta.$$

Thus, irrespective of the direction θ one moves away from the critical point, the function f *always* increases. Thus there is a local minimum at the critical point.

(iii) Suppose $rt - s^2 < 0$. We wish to show the critical point is a saddle point. Our proof below ASSUMES $t \neq 0$. The case where $t = 0$ is a straightforward alteration and left as an exercise.

Part I. We first of all wish to show, for $r,\ t,\ s$ fixed and satisfying the aforementioned constraint that

$$\frac{d^2f}{dn^2}(a,b) = r\left(\cos\theta + \frac{s}{r}\sin\theta\right)^2 + \frac{1}{r}\left(rt - s^2\right)\sin^2\theta = 0 \tag{5.1}$$

has four real roots for $\theta \in [0,\ 2\pi)$. In Part II, we show why this is sufficient to enforce a saddle point.

- Equation (5.1) does not have a root for θ satisfying $\cos\theta = 0$.
  - Suppose $\cos\theta = 0$. Then $\sin^2\theta = 1$ and hence $d^2f/dn^2 = t \neq 0$ (by above assumption).

- Equation (5.1) has four roots.
  - At a root, we have $\cos\theta \neq 0$, and thus
$$\sec^2\theta \frac{d^2 f}{dn^2}(a,b) = r + 2s\tan\theta + t\tan^2\theta = 0$$
    which yields
$$\theta = \tan^{-1}\left(\frac{-s \pm \sqrt{s^2 - rt}}{2t}\right).$$
    Noting $s^2 - rt > 0$, this gives four real roots for θ within $[0, 2\pi)$.

Part II. We have that $d^2f/dn^2 = 0$ for four values of $\theta \in [0, 2\pi)$. Thus, on this domain, there are two non-adjacent intervals of θ where $d^2f/dn^2 > 0$; moving away from the critical in these directions leads to an increase in f. Similarly, there are two non-adjacent intervals of θ where $d^2f/dn^2 < 0$; moving away from the critical in these directions leads to a decrease in f. Thus we have a saddle point. ■

Remark The above is somewhat inelegant. The use of Taylor's expansion for a function of two variables, and the fact a symmetric matrix can be diagonalised via an orthogonal transformation, readily leads to a proof of the above; we do not cover these topics in this course.

Example 90 *For each of the functions below find the critical points and classify them in terms of maxima, minima or saddle points:*

(i) $f(x, y) = x + 2x^2y^2 - x^2 + y^2$

(ii) $g(x, y) = \cos(x + y)\sin x, \quad$ for $x, y \in (0, \pi)$.

Solution.

(i) At a critical point we must have
$$f_x = 1 + 4xy^2 - 2x = 0, \qquad f_y = 4x^2y + 2y = y(4x^2 + 2) = 0.$$
For $f_y = 0$ we must have $y = 0$, whence $x = 1/2$ from the left of the above. Thus we have a single critical point at $(x, y) = (1/2, 0)$. Noting that
$$f_{xx} = 4y^2 - 2, \qquad f_{xy} = f_{yx} = 8xy, \qquad f_{yy} = 4x^2 + 2$$
and using the above definitions of $r,\ s,\ t$ we have
$$r = -2, \quad s = 0, \quad t = 3, \quad rt - s^2 = -6.$$
Hence the critical point at $(x, y) = (0, 1/2)$ is a saddle.

(ii) At a critical point we must have

■

6. PARTIAL DIFFERENTIAL EQUATIONS

6.1 Laplace's Equation

We have already mentioned **Laplace's equation** in earlier examples. In Cartesian co-ordinates it reads as

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} &= 0 \qquad \text{(in the plane);} \\ \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} &= 0 \qquad \text{(in three dimensions).} \end{aligned}$$

In Example 30 we showed how the equation remained the same when we changed from Cartesian to parabolic co-ordinates, and in Example 31 we saw that Laplace's equation reads as

$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} = 0$$

in planar polar co-ordinates. A function which satisfies Laplace's equation is known as **harmonic**.

Laplace's equation is often abbreviated to

$$\nabla^2 f = 0$$

and the differential operator ∇^2 is known as the **Laplacian.**

Laplace's equation appears in many physical situations. For example, $\nabla^2 f = 0$ holds true for:

- the gravitational potential in a region containing no matter;
- the electrostatic potential in a charge-free region;
- the steady-state temperature in a region containing no source of heat;
- incompressible fluid flows whenever viscosity is negligible and there are no vortices, sources or sinks. Then the equations of motion for the fluid can be recast into Laplace's equation for a velocity potential.

Example 91 *Let* $z = x + iy$ *and let* $f(z) = z^n$. *If we write* $u = \operatorname{Re} f$ *and* $v = \operatorname{Im} f$ *so that*

$$u(x,y) + iv(x,y) = (x+iy)^n, \tag{6.1}$$

show that $u(x,y)$ *and* $v(x,y)$ *are both harmonic functions.*

Solution. If we differentiate (6.1) twice with respect to x, and also twice with respect to y, we get

$$\begin{aligned} u_{xx} + iv_{xx} &= n(n-1)(x+iy)^{n-2}; \\ u_{yy} + iv_{yy} &= n(n-1)i^2(x+iy)^{n-2} = -n(n-1)(x+iy)^{n-2}. \end{aligned}$$

Hence, adding the previous two equations,

$$(u_{xx} + u_{yy}) + i(v_{xx} + v_{yy}) = 0.$$

Comparing real and imaginary parts we see that u and v are both harmonic. ■

Exercise 92 *Show that the real and imaginary parts of*

$$\begin{aligned} e^z &= e^{x+iy} = e^x \cos x + ie^x \sin y; \\ \frac{1}{z} &= \frac{1}{x+iy} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2} \end{aligned}$$

are also harmonic.

Separable Solutions We will first consider **separable solutions** of Laplace's equation in Cartesian co-ordinates. That is we shall look for solutions of the form

$$f(x,y) = X(x)\,Y(y) \quad \text{for} \quad \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

If we substitute $f(x,y) = X(x)\,Y(y)$ into Laplace's equation we get that

$$X''(x)\,Y(y) + X(x)\,Y''(y) = 0.$$

If we rearrange this to

$$\frac{X''(x)}{X(x)} = -\frac{Y''(y)}{Y(y)}. \tag{6.2}$$

The quantity, let's call it $c(x,y)$, in (6.2) is both a function solely of x (from the LHS) and a function solely of y (from the RHS). It follows that for any $(x_1,y_1), (x_2,y_2)$

$$\begin{aligned} c(x_1,y_1) &= c(x_1,y_2) && [\text{as } c \text{ only depends on } x] \\ &= c(x_2,y_2) && [\text{as } c \text{ only depends on } y]. \end{aligned}$$

That is, c is constant! So we have

$$\begin{aligned} \frac{X''(x)}{X(x)} &= c, \\ \frac{Y''(y)}{Y(y)} &= -c. \end{aligned}$$

- If $c = 0$ we see that

  $$X(x) = Ax + B, \qquad Y(y) = Cy + D$$

  for constant A, B, C, D.
- If $c > 0$ we see that

  $$\begin{aligned} X(x) &= A\exp\left(\sqrt{c}x\right) + B\exp\left(-\sqrt{c}x\right) \\ Y(y) &= C\cos\left(\sqrt{c}y\right) + D\sin\left(\sqrt{c}y\right) \end{aligned}$$

  for constants A, B, C, D.
- If $c < 0$ we have a situation similar to the $c > 0$ case with x and y swapped.

Example 93 *Find the separable solutions $R(r)\Theta(\theta)$ to Laplace's equation in planar polar co-ordinates.*

Solution. Putting $f(r,\theta) = R(r)\Theta(\theta)$ into

$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} = 0$$

we get

$$R''(r)\Theta(\theta) + \frac{1}{r}R'(r)\Theta(\theta) + \frac{1}{r^2}R(r)\Theta''(\theta) = 0.$$

This rearranges to

$$\frac{r^2 R''}{R} + \frac{rR'}{R} = -\frac{\Theta''}{\Theta}$$

both sides of which must be constant, as previously shown. If we call this constant c then we have ...

- If $c = k^2 > 0$ then we have that

- If $c = 0$ then we have

- If $c = -k^2 < 0$ then we have

■

Remark 94 *Note that in the cases $c = k^2 > 0$ the solutions*

$$r^k \cos k\theta, \qquad r^k \sin k\theta, \qquad r^{-k} \cos k\theta, \qquad r^{-k} \sin k\theta$$

are the real and imaginary parts of z^k, then of z^{-k}, respectively.

6.2 Dirichlet's Problem

Solving Laplace's equation

$$\nabla^2 f = 0$$

in a region R, under a boundary condition

$$f(x) = g(x) \text{ for } x \in \partial R$$

on the boundary ∂R of the region R, is known as *Dirichlet's Problem*. This is named after the nineteenth century German mathematician Peter Gustav Dirichlet (1805-1859). Under fairly mild conditions on the shape of the boundary ∂R and the function $g(x)$ it can be shown that there exists a unique solution to Dirichlet's problem.

It is beyond the scope of the course to treat any aspect of this problem in a general setting but we will look at the case of a rectangular and circular region.

Example 95 Part I *Find the most general sum of separable solutions defined on the rectangle*

$$R = \{(x, y) : 0 \leqslant x \leqslant a,\ 0 \leqslant y \leqslant b\}$$

which satisfies Laplace's equation in the interior of R,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0, \qquad 0 < x < a,\ 0 < y < b$$

and which satisfies

$$\begin{aligned} f(0, y) &= 0, \qquad 0 < y < b, &(6.3)\\ f(x, 0) &= 0, \qquad 0 < x < a, &(6.4)\\ f(x, b) &= 0, \qquad 0 < x < a. &(6.5) \end{aligned}$$

Part II *Hence find the solution to*

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0, \qquad 0 < x < a,\ 0 < y < b$$

given the boundary conditions

$$\begin{aligned}
f(0,y) &= 0, \qquad 0<y<b,\\
f(x,0) &= 0, \qquad 0<x<a,\\
f(x,b) &= 0, \qquad 0<x<a,\\
f(a,y) &= g(y) \overset{def}{=} 2\sin\left(\frac{\pi y}{b}\right)+\sin\left(\frac{3\pi y}{b}\right), \qquad 0<y<b.
\end{aligned}$$

Solution. Part I We need to find the most general sum of separable solutions and begin by considering the separable solutions we met earlier. These were of the form

$$f(x,y) = \left\{A\exp\left(\sqrt{c}x\right)+B\exp\left(-\sqrt{c}x\right)\right\}\left\{C\cos\left(\sqrt{c}y\right)+D\sin\left(\sqrt{c}y\right)\right\}, \qquad c>0; \tag{6.6}$$

$$f(x,y) = \left\{A\cos\left(\sqrt{c}x\right)+B\sin\left(\sqrt{c}x\right)\right\}\left\{C\exp\left(\sqrt{c}y\right)+D\exp\left(-\sqrt{c}y\right)\right\}, \qquad c>0; \tag{6.7}$$

$$f(x,y) = (Ax+B)(Cy+D). \tag{6.8}$$

If we consider which of these can satisfy

$$f(x,0)=f(x,b)=0, \qquad 0<x<a,$$

then we see:

in (6.6) this means that

in (6.7) this means that

in (6.8) this means that

Hence the non-zero separable solutions are of the form

for some integer n. [The constant D is now no longer necessary.]

If we also require that $f(0,y)=0$ for all $y\in(0,b)$ then we have

It follows that the most general sum of separable solutions satisfying Laplace's equation in the rectangle R and the conditions (6.3)-(6.5) is

(6.9)

Part II We additionally need to impose the fourth boundary condition that

$$f(a,y)=g(y)\overset{def}{=}2\sin\left(\frac{\pi y}{b}\right)+\sin\left(\frac{3\pi y}{b}\right),\qquad 0<y<b,$$

Comparing coefficients with the expression (6.9) we see that

Hence

$$f(x,y)=\frac{2}{\sinh(\pi a/b)}\sin\left(\frac{\pi x}{a}\right)\sin\left(\frac{\pi y}{b}\right)+\frac{1}{\sinh(3\pi a/b)}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{3\pi y}{b}\right).$$

■

Remark 96 *We will briefly comment in the next section on how to deal with a more general function $g(y)$ in the boundary condition. However, it would also be reasonable to remark that the given boundary condition has three of its sides set to 0. This, though, is not as limiting as a restriction as one might first think. Given a general boundary condition on the four sides of the rectangle, we can treat each side separately as above to produce four functions that match the boundary condition on a single side. Their sum is a harmonic function which meets the required boundary condition on all four sides.*

Example 97 *Let $f(r,\theta)$ be a function, defined on the disc $R = \{(r,\theta) : r \leqslant 1\}$ which satisfies Laplace's equation in the interior of R,*

$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} = 0, \qquad r < 1$$

and satisfies the boundary condition $f(1,\theta) = g(\theta)$, $\qquad 0 \leqslant \theta < 2\pi$, for some function g.

Solution. We'll begin by considering the separable solutions in r and θ we met earlier. These were of the form

$$\begin{aligned} f(r,\theta) &= \left\{Ar^k + Br^{-k}\right\}\{C\cos k\theta + D\sin k\theta\} \qquad k > 0; \\ f(r,\theta) &= \{A\sin(k\ln r) + B\cos(k\ln r)\}\left\{Ce^{k\theta} + De^{-k\theta}\right\}, \qquad k > 0; \\ f(r,\theta) &= \{A\ln r + B\}\{Ct + D\}. \end{aligned}$$

Note that the only solutions which are defined at $r = 0$ (which lies in R) are those of the form

$$f(r,\theta) = r^k (C\cos k\theta + D\sin k\theta),$$

where k is a non-negative integer (so that it is continuous on the $\theta = 0, 2\pi$ borderline.)

Again it is the case that any sum of these

$$f(r,\theta) = \sum_{k=0}^{\infty} r^k (C_k\cos k\theta + D_k\sin k\theta)$$

is also a solution of the Laplace's equation. If the function $g(\theta)$ is of the form

$$g(\theta) = \cos^2\theta + \sin\theta$$

then we can rewrite this as

$$g(\theta) = \frac{1}{2} + \frac{1}{2}\cos 2\theta + \sin\theta,$$

and comparing coefficients we see that

$$C_0 = \frac{1}{2}, \qquad C_2 = \frac{1}{2}, \qquad D_1 = 1$$

with all other coefficients zero. Hence

$$f(r,\theta) = \frac{1}{2} + \frac{1}{2}r^2\cos 2\theta + r\sin\theta.$$

■

6.3 Fourier Analysis

Recall when solving Dirichlet's Problem in the rectangle we produced solutions

$$f(x,y) = \sum_{n=1}^{\infty} \alpha_n \sinh\left(\frac{n\pi x}{b}\right) \sin\left(\frac{n\pi y}{b}\right)$$

which met the boundary conditions on the three sides where the function was zero.

We could meet the additional boundary condition on $x = a$ easily when the function there, $g(y)$, was a finite linear combination of the functions

$$\sin\left(\frac{n\pi y}{b}\right),$$

as we could compare coefficients naturally.

Likewise with Dirichlet's Problem in the circle we needed to solve the boundary condition

$$\sum_{k=0}^{\infty} (C_k \cos k\theta + D_k \sin k\theta) = g(\theta)$$

on $r = 1$. Again we can find the coefficients easily if the function $g(\theta)$ is a finite linear combination of functions

$$\cos k\theta \quad \text{and} \quad \sin k\theta.$$

— —

The more general problem of writing a function $F(t)$ as an infinite sum of function $\cos kt$ and $\sin kt$ is the subject of **Fourier analysis**. Compare this with the situation where $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are orthonormal vectors (mutually perpendicular and of unit length) and we have

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i.$$

In order to work out each α_i we look can take the component in the $\mathbf{e}_i$ direction. If we dot both sides of the equation by $\mathbf{e}_i$ then we find

$$\mathbf{v} \bullet \mathbf{e}_i = \alpha_i.$$

Similarly when faced with the equation

$$F(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt), \qquad t \in [0, 2\pi] \tag{6.10}$$

we are looking for a way to extract each coefficient a_k or b_k, somehow taking a component in the "direction" $\cos kt$ or $\sin kt$.

In Fourier analysis this is done by noting for $k_1, k_2 > 0$,

$$\begin{aligned}
\int_0^{2\pi} \sin(k_1 t) \sin(k_2 t) \; \mathrm{d}t &= \\
&= \begin{cases} 0 & \text{if } k_1 \neq k_2 \\ \pi & \text{if } k_1 = k_2 \end{cases} \\
\int_0^{2\pi} \cos(k_1 t) \cos(k_2 t) \; \mathrm{d}t &= \begin{cases} 0 & \text{if } k_1 \neq k_2 \\ \pi & \text{if } k_1 = k_2 \end{cases} \\
\int_0^{2\pi} \sin(k_1 t) \cos(k_2 t) \; \mathrm{d}t &= 0 \text{ for all } k_1, k_2.
\end{aligned}$$

So "dotting" (6.10) with $\cos kt$, by which is meant "multiplying both sides by $\cos kt$ and integrating over $[0, 2\pi]$" we find

$$\int_0^{2\pi} F(t) \cos kt \; \mathrm{d}t \quad =$$

Thus, the "dotting" operation eliminates all other coefficients as required.

So the formulae for the Fourier coefficients are.

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_0^{2\pi} F(t)\cos kt\ \mathrm{d}t \qquad (k>0) \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} F(t)\sin kt\ \mathrm{d}t \qquad (k>0) \\ a_0 &= \frac{1}{\pi}\int_0^{2\pi} F(t)\ \mathrm{d}t. \end{aligned}$$

provided all the above calculations and the interchange of integration and summation can be justified and issues of convergence overcome! Fourier series and their convergence will be dealt with in detail in next term's Fourier Series course.

Example 98 *Find the Fourier coefficients of*

$$f(x) = x, \qquad x \in [0, 2\pi].$$

Consider the Fourier series' convergence at $x = 0, \pi/2, \pi, 3\pi/2, 2\pi$ *and at* $x + 2\pi$ *compared with* x.

Solution. The Fourier coefficients are

$$\begin{aligned} a_0 &= \frac{1}{\pi}\int_0^{2\pi} x\,\mathrm{d}x = \left[\frac{x^2}{2\pi}\right]_0^{2\pi} = 2\pi; \\ a_k &= \frac{1}{\pi}\int_0^{2\pi} x\cos kx\,\mathrm{d}x = \frac{1}{\pi}\left[\frac{x\sin kx}{k} + \frac{\cos kx}{k^2}\right]_0^{2\pi} = 0; \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} x\sin kx\,\mathrm{d}x = \frac{1}{\pi}\left[\frac{-x\cos kx}{k} + \frac{\sin kx}{k^2}\right]_0^{2\pi} = \frac{-2}{k}. \end{aligned}$$

So the Fourier series of $f(x) = x$ is

$$F(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k\cos kx + b_k\sin kx) = \pi - 2\sum_{k=1}^{\infty}\frac{\sin kx}{k}$$

Consider the convergence of this series.

- The first thing to note is that $F(x+2\pi) = F(x)$ for all x because F is a sum of sines. So certainly F does not agree with f everywhere.
- If we look at F at $x = 0, \pi, 2\pi$ we see that F equals π at each of these points.
- At $x = \pi/2$ and $x = 3\pi/2$ we have

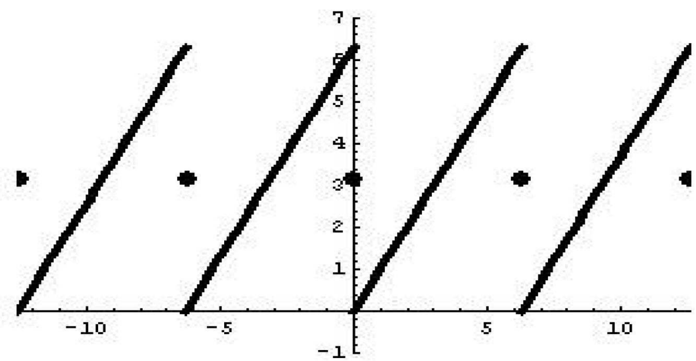
$$\begin{aligned} F\left(\frac{\pi}{2}\right) &= \pi - 2\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots\right) = \pi - 2\left(\frac{\pi}{4}\right) = \frac{\pi}{2}; \\ F\left(\frac{3\pi}{2}\right) &= \pi - 2\left(-1 + \frac{1}{3} - \frac{1}{5} + \frac{1}{7} - \cdots\right) = \pi - 2\left(\frac{-\pi}{4}\right) = \frac{3\pi}{2}. \end{aligned}$$

  So at least the Fourier series converges correctly at $x = \pi/2, \pi, 3\pi/2$. The infinite series

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}$$

  was first proved by Leibniz.
- In fact, the series $F(x)$ converges to x on the interval $(0, 2\pi)$. At multiples of 2π the series converges to π — more generally a Fourier series converges at multiples of 2π to the average of $f(0)$ and $f(2\pi)$, and will do similarly at any discontinuity in the interval $(0, 2\pi)$.

The graph of $F(x)$ is below:



Example 99 *By considering the Fourier series of x at $\pi/4$ find the sum of the infinite series*

$$S = 1 + \frac{1}{3} - \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \frac{1}{11} - \frac{1}{13} - \frac{1}{15} + \cdots$$

Solution. If we put $x = \pi/4$ into $F(x)$ we get

$$\frac{\pi}{4} = F\left(\frac{\pi}{4}\right) =$$

Thus $S = \pi/\left(2\sqrt{2}\right)$. ■

6.4 Poisson's Equation

Poisson's equation is the inhomogeneous Laplace equation

$$\nabla^2 f = g.$$

It commonly appears in gravitational theory: Laplace's equation $\nabla^2\phi = 0$ dictates how the *gravitational potential* ϕ behaves in the absence of any matter, but, more generally, when there is matter present, distributed with density function $\rho(x, y, z)$, then Poisson's equation

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

describes the potential's behaviour. Poisson's equation occurs in numerous other physical scenarios, such as electrostatics in the presence of charge distributions.

Whilst we will not be specifically interested in gravity, electrostatics or other physical applications in this course, solving Poisson's equation raises the same issues that will again appear in later courses.

Example 100 *Find the circularly symmetric solutions of Poisson's equation in the plane $\nabla^2 f = g$ where*

$$g(r) = \begin{cases} r & \text{for } r < a \\ 0 & \text{for } r > a \end{cases}$$

Solution. Recall that the Laplacian is given by

$$\nabla^2 f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}$$

in planar polar co-ordinates. We are only interested in circularly symmetric functions of the form $f(r)$; so for $r < a$ we need to solve

$$f''(r) + r^{-1} f'(r) = r.$$

If we solve the equation similarly in the region $r > a$ then we find

There are several things to note though. Firstly if this function is to be defined at $r = 0$ it must be the case that $A = 0$. But also there are issues of continuity and smoothness at the $r = a$ border. For our solution f to be continuous at $r = a$ we need

$$a^3/9 + B = C \ln a + D$$

and for it to be smooth at $r = a$ (that is $\mathrm{d}f/\mathrm{dr}$ agrees on both sides of $r = a$) we need $a^2/3 = C/a$. Hence

$$A = 0, \qquad C = a^3/3, \qquad D = \left(a^3/9\right)(1 - 3\ln a) + B.$$

So the most general circularly symmetric solution of this Poisson equation is

$$f(r) = \begin{cases} \frac{r^3}{9} + B & \text{for } r \leqslant a, \\ \frac{a^3}{9}\ln r + \frac{a^3}{9}(1 - 3\ln a) + B & \text{for } r \geqslant a. \end{cases}$$

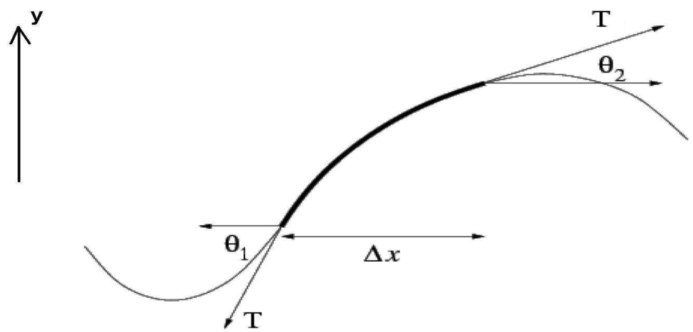
To specify the solution uniquely then we might, for example, require a further condition such as $f(0) = 0$ which would give $B = 0$. ■

6.5 The Wave Equation

Derivation We will consider the vibrations of an elastic uniform string under certain simplifying assumptions:-

- the string undergoes small vibrations: by this we mean that second order terms, such as y^2, y_x^2, θ_1^2 and θ_2^2 in the analysis below will be considered to be negligible;
- the vibrations are entirely *transverse*, so that a point at distance x_0 along the string remains on the line $x = x_0$ throughout the motion;
- the string is at constant tension T and the density of the string is uniformly ρ;
- the effects of gravity and air resistance are negligible compared with the tension in the string.

Consider the vibrations of a small section of the string from x to $x + \Delta x$. We shall denote the angles the string makes with horizontal at x and $x + \Delta x$ by θ_1 and θ_2 respectively, as shown in the diagram below.



I. Taking components of the forces in the y-direction and invoking Newton's second law (Force = Mass × Acceleration) we get

where x_0 is the x-co-ordinate of the centre of mass of the segment, $\rho\Delta x$ is the segment's mass and y_{tt} its vertical acceleration. Note that x_0 will lie between x and $x + \Delta x$ as the entirety of the segment of string lies in this range.

There is no gravity term as the weight of the string is considered negligible compared with the tension involved.

II At the level of approximation we are considering (i.e. neglecting second order terms), we have

$$\sin\theta_1 = \theta_1 - \frac{\theta_1^3}{6} + \ldots \quad \sim \theta_1 \sim \tan\theta_1 = \theta_1 + \frac{\theta_1^3}{3} + \ldots$$

and similarly for θ_2. Hence at the level of approximation we are considering

$$\frac{\tan\theta_2 - \tan\theta_1}{\Delta x} = \frac{\rho}{T}\frac{\partial^2 y}{\partial t^2}(x_0, t)\,.$$

III Now

$$\begin{aligned}\tan\theta_1 &= \frac{\partial y}{\partial x}(x,t)\,,\\ \tan\theta_2 &= \frac{\partial y}{\partial x}(x+\Delta x,t) =\end{aligned}$$

where the final expression can be derived by a Taylor expansion (assuming y_{xx} is continuous and differentiable; see other courses this year for details of Taylor expansions). Thus

$$\frac{\partial^2 y}{\partial x^2}(x,t) - \frac{\rho}{T}\frac{\partial^2 y}{\partial t^2}(x_0,t) = O(\Delta x).$$

Letting Δx tend to 0, whereupon $x_0 \to x$, yields

$$\frac{\partial^2 y}{\partial x^2}(x,t) = \frac{\rho}{T}\frac{\partial^2 y}{\partial t^2}(x,t)\,.$$

Note that ρ/T has units $(\mathrm{kg/m})/(\mathrm{kgms^{-2}}) = (\mathrm{s/m})^2$ so that $c = \sqrt{T/\rho}$ has the units of velocity. The wave-equation then reads

$$c^2\frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}.$$

Proposition 101 *(D'Alembert 1746) The general solution of the wave equation*

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}$$

is

$$y(x,t) = f(x-ct) + g(x+ct).$$

Proof. *We introduce two new variables*

$$\zeta = x - ct \quad \text{and} \quad \eta = x + ct.$$

Then by Theorem 29, that is the chain rule, we have

$$\begin{aligned} y_{xx} &= y_\zeta \zeta_{xx} + y_\eta \eta_{xx} + y_{\zeta\zeta}(\zeta_x)^2 + 2y_{\zeta\eta}\zeta_x\eta_x + y_{\eta\eta}(\eta_x)^2 \\ &= y_\zeta \times 0 + y_\eta \times 0 + y_{\zeta\zeta} \times (1)^2 + 2y_{\zeta\eta} \times 1 \times 1 + y_{\eta\eta} \times (1)^2 \\ &= y_{\zeta\zeta} + 2y_{\zeta\eta} + y_{\eta\eta} \end{aligned}$$

and

$$\begin{aligned} y_{tt} &= y_\zeta \zeta_{tt} + y_\eta \eta_{tt} + y_{\zeta\zeta}(\zeta_t)^2 + 2y_{\zeta\eta}\zeta_t\eta_t + y_{\eta\eta}(\eta_t)^2 \\ &= y_\zeta \times 0 + y_\eta \times 0 + y_{\zeta\zeta} \times (-c)^2 + 2y_{\zeta\eta} \times (-c) \times c + y_{\eta\eta} \times (c)^2 \\ &= c^2 (y_{\zeta\zeta} - 2y_{\zeta\eta} + y_{\eta\eta}) \end{aligned}$$

Hence $c^2 y_{xx} = y_{tt}$ *if*

$$c^2 (y_{\zeta\zeta} + 2y_{\zeta\eta} + y_{\eta\eta}) = c^2 (y_{\zeta\zeta} - 2y_{\zeta\eta} + y_{\eta\eta})$$

so that $y_{\zeta\eta} = 0$. *We know, from Example 12(i), that the general solution of this is*

$$y = f(\zeta) + g(\eta) = f(x-ct) + g(x+ct).$$

■

Remark 102 *Consider a solution of the form $y(x,t) = f(x - ct)$, with $g = 0$. Note that this solution at time $t + T$ resembles the solution at time t, but translated to the right by cT. This is a wave moving to the right with speed c and likewise a wave of the form $g(x + ct)$ is one which is moving to the left at speed c.*

Example 103 *Find the solution to the wave equation when*

$$y(x,0) = \begin{cases} 1 - |x| & -1 < x < 1 \\ 0 & \textit{otherwise} \end{cases}, \qquad \frac{\partial y}{\partial t}(x,0) = 0.$$

Sketch your solutions at $ct = 0, \frac{1}{2}, 1, \frac{3}{2}$.

Solution. We know that the solution has the form $y(x,t) = f(x - ct) + g(x + ct)$ and so the initial conditions give that

$$\begin{aligned} f(x) + g(x) &= \begin{cases} 1 - |x| & -1 < x < 1, \\ 0 & \text{otherwise}, \end{cases} \\ -cf'(x) + cg'(x) &= 0. \end{aligned}$$

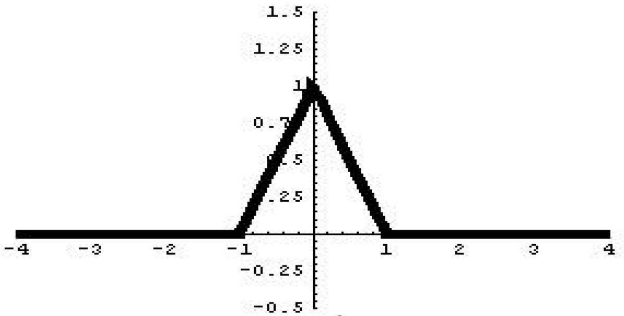
Integrating the second equation we see that $f(x) = g(x) + K$ where K is a constant. We can, without any loss of generality take K to be 0. [Keeping K in the calculation would yield to the same solution with slightly different f and g.] So $f = g$ and we have

$$y(x,t) = f(x - ct) + f(x + ct)$$

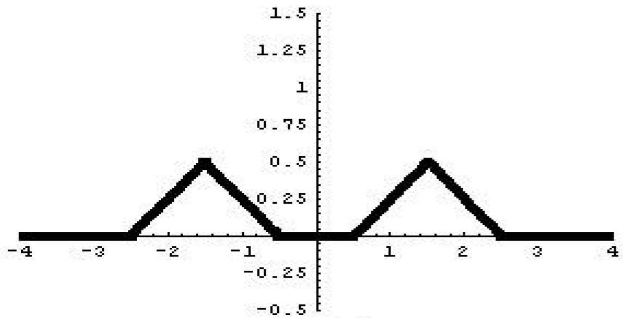
where

$$f(u) = \begin{cases} \frac{1}{2}(1 - |u|) & -1 < u < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Sketches of the solution at different t are given below.



$t = 0$



$ct = 1.5$

■

Example 104 *Find all the separable solutions*

$$y(x,t) = X(x)\,T(t) \tag{6.11}$$

of the wave equation $y_{tt} = c^2 y_{xx}$ *which satisfies the boundary conditions*

$$y(0,t) = y(L,t) = 0 \quad \textit{for all } t.$$

Solution. Substituting $y(x,t) = X(x)\,T(t)$ into Equation (6.11) we get

$$X(x)\,T''(t) = c^2 X''(x)\,T(t) \Longrightarrow \frac{X''(x)}{X(x)} = \frac{T''(t)}{c^2 T(t)} = k\,(\text{constant}).$$

We then have

$$X(x) =$$

depending on the sign of k.

As in Example 95 the only one of these solutions which can meet the boundary conditions $X(0) = X(L) = 0$ without being everywhere zero is the third of these.

For

$$X(t) = E\cos\left(\sqrt{-k}x\right) + F\sin\left(\sqrt{-k}x\right)$$

to satisfy $X(0) = 0$ then we must have $E = 0$.

In order to $X(L) = 0$ without having $F = 0$ it must be the case that

$$\sqrt{-k} =$$

for some integer n. So the separable solutions are of the form

$$y(x,t) = X(x)\,T(t) = \sin\left(\frac{n\pi x}{L}\right)\left\{\alpha_n \cos\left(\frac{n\pi ct}{L}\right) + \beta_n \sin\left(\frac{n\pi ct}{L}\right)\right\}.$$

■

Remark 105 *The solutions of the form*

$$y(x,t) = A \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{n\pi ct}{L}\right)$$

*are known as the **natural** or **normal modes**; when $n = 1$ this is known as the **fundamental mode**. The quantities $n\pi c/L$ are called the **natural frequencies** of the string.*

Remark 106 *Using Fourier analysis as in Section 6.3 it is possible to write down any solution of the wave equation with these boundary conditions in terms of the separable solutions. To specify uniquely the solution initial conditions describing $y(x,0)$ and $y_t(x,0)$ can also be given.*

Example 107 *A string with displacement $y(x,t)$, where $a \leqslant x \leqslant b$, is fixed at each end so that $y(a,t) = 0 = y(b,t)$ for all time t. Show that the string's energy*

$$E(t) = \int_a^b \left\{ \frac{1}{2} T \left(\frac{\partial y}{\partial x}\right)^2 + \frac{1}{2}\rho \left(\frac{\partial y}{\partial t}\right)^2 \right\} dx$$

is constant throughout the motion. The first term of the energy formula is the tensile energy and the second term represents the string's kinetic energy.

Solution. To show E is constant, we will differentiate it and make use of the wave equation

$$y_{tt} = \frac{T}{\rho} y_{xx}$$

to get

$$\begin{aligned}
E'(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_a^b \left\{\frac{1}{2}T\left(\frac{\partial y}{\partial x}\right)^2 + \frac{1}{2}\rho\left(\frac{\partial y}{\partial t}\right)^2\right\}\mathrm{d}x \\
&= \int_a^b \frac{\partial}{\partial t}\left\{\frac{1}{2}T\left(\frac{\partial y}{\partial x}\right)^2 + \frac{1}{2}\rho\left(\frac{\partial y}{\partial t}\right)^2\right\}\mathrm{d}x \\
&= \int_a^b (T\,y_x y_{xt} + \rho\,y_t y_{tt})\,\mathrm{d}x \\
&= \int_a^b (T y_x y_{xt} + T y_t y_{xx})\,\mathrm{d}x \qquad \text{[by the wave equation]} \\
&= T\int_a^b \left\{\frac{\mathrm{d}}{\mathrm{d}x}(y_x y_t)\right\}\mathrm{d}x \\
&= T\left(y_x(b,t)\,y_t(b,t) - y_x(a,t)\,y_t(a,t)\right) = 0.
\end{aligned}$$

Hence $E(t)$ is constant. ■

6.6 The Heat Equation

We will consider here the problem of modelling heat flow in a one-dimensional uniform bar. If two nearby points on the rod, separately by a small distance ϵ, are at temperatures T_1 on the left and T_2 on the right, then the heat flow *from left to right* between these points is proportional to the temperature difference and inversely proportional to the distance. That is

$$\text{Amount of heat per unit time } = \kappa \frac{T_1 - T_2}{\epsilon} \tag{6.12}$$

where the constant of proportionality κ is called the thermal conductivity and which (we assume) depends only on the material that makes up the rod. Equation (6.12) is **Fourier's law** of heat conduction. Note that this heat is "signed" in the sense that it is negative if $T_2 > T_1$ and the heat is actually flowing from right to left.

Consider a small section of the rod between $x = x_0$ and $x = x_0 + \delta x$. We'll consider the *flux* of heat into this section. The rate of heat transfer from left-to-right past the point $x = x_0$ equals

$$\lim_{\epsilon \to 0} k \frac{T\left(x_0 - \epsilon/2, t\right) - T\left(x_0 + \epsilon/2\right)}{\epsilon} = -k \frac{\partial T}{\partial x}\left(x_0, t\right)$$

and similarly the rate at which heat flows from left to right past the point $x = x_0 + \delta x$ equals

$$-\kappa \frac{\partial T}{\partial x}\left(x_0 + \delta x, t\right)$$

so that the amount of heat entering the section, in a short time δt equals

$$\kappa \left[\frac{\partial T}{\partial x}\left(x_0 + \delta x, t\right) - \frac{\partial T}{\partial x}\left(x_0, t\right) \right] \delta t.$$

Now the average change of temperature δT in the small section of the rod is proportional to the amount of heat introduced (given above) and inversely proportional to the mass $\rho\,\delta x$ of the section. The constant of proportionality is c^{-1} where c is called the specific heat of the material. Hence

$$\delta T =$$

If we rearrange this to

$$\frac{\delta T}{\delta t} = \frac{\kappa}{\rho c}\left[\frac{\partial T}{\partial x}(x_0 + \delta x, t) - \frac{\partial T}{\partial x}(x_0, t)\right]\Big/\,\delta x$$

and let δx and δt tend to zero then we find

$$\frac{\partial T}{\partial t} = \frac{\kappa}{\rho c}\frac{\partial^2 T}{\partial x^2}. \tag{6.13}$$

Equation (6.13) is known as the **heat equation** or **diffusion equation**. The constant $\kappa/\rho c$ is often written as α^2 and is called the **thermal diffusivity** of the rod so that the heat equation reads

$$\frac{\partial T}{\partial t} = \alpha^2\frac{\partial^2 T}{\partial x^2}.$$

More generally, in three dimensions, the heat equation reads

$$\frac{\partial T}{\partial t} = \alpha^2\nabla^2 T.$$

Note that the steady-state solutions of the heat equation, i.e. those solutions that don't depend on time, are the solutions of Laplace's equation

$$\nabla^2 T = 0.$$

When modelling the heat flow in such a rod $0 \leqslant x \leqslant L$ there are two natural boundary conditions which might arise.

- An end of the rod may be kept at a particular temperature. For example we might have

$$T\,(L, t) = T_0 \quad \text{for all } t.$$

- An end of the rod may be insulated so that no heat is lost through that end. For example, if the end at $x = 0$ were insulated it would mean
$$\frac{\partial T}{\partial x}(0,t) = 0 \text{ for all } t.$$

Consider the example below. To solve it completely we will have to calculate the separable solutions of the heat equation and calculate another Fourier series.

Example 108 *The temperature T in a rod $0 \leqslant x \leqslant L$ is initially T_0. The ends of the rod are kept at temperature $T = 0$; that is for all t we have*
$$T(0,t) = T(L,t) = 0. \tag{6.14}$$
Determine the temperature $T(x,t)$ in the rod at time t and at position x.

Solution. We begin by determining the separable solutions to the heat equation that satisfy the boundary conditions (6.14). If we substitute
$$T(x,t) = A(x)B(t)$$
into the heat equation
$$\frac{\partial T}{\partial t} = \alpha^2 \frac{\partial^2 T}{\partial x^2}$$
we find
$$A(x)B'(t) = \alpha^2 A''(x)B(t)$$
and separating the variables we get
$$\frac{A''(x)}{A(x)} = \frac{B'(t)}{\alpha^2 B(t)} = k \text{ (constant)},$$
as only constants are functions of x alone and also of t alone. So we have, depending on the sign of k
$$A(x) = \{$$
As we have seen before, it is only the third of these, when $k < 0$ that yields any non-zero solutions meeting the boundary conditions
$$A(0) = A(L) = 0.$$
In this case we have solutions of the form
$$A(x) = Q\sin\left(\frac{n\pi x}{L}\right)$$

where $\sqrt{-k} = n\pi/L$. Substituting this back into the equation

$$B'(t) = k\alpha^2 B(t)$$

we get separable solutions of the form

$$T(x,t) = Q_n \sin\left(\frac{n\pi x}{L}\right) \exp\left\{\frac{-n^2\pi^2\alpha^2 t}{L^2}\right\}.$$

A general solution of the heat equation, satisfying the boundary conditions (6.14) is of the form

$$T(x,t) = \sum_{n=1}^{\infty} Q_n \sin\left(\frac{n\pi x}{L}\right) \exp\left\{\frac{-n^2\pi^2\alpha^2 t}{L^2}\right\}.$$

If this is solution is going to meet the initial condition that $T(x,0) = T_0$ in the rod then we must have

$$T_0 = \sum_{n=1}^{\infty} Q_n \sin\left(\frac{n\pi x}{L}\right). \tag{6.15}$$

Recall that a crucial idea behind Fourier analysis is that

$$\int_0^L \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \mathrm{d}x = \begin{cases} 0 & \text{if } n \neq m, \\ L/2 & \text{if } n = m, \end{cases}$$

though our earlier examples in Section 6.3 we on an interval of length 2π rather than the more general L here. Hence

$$\begin{aligned} Q_n &= \frac{2T_0}{L} \int_0^L \sin\left(\frac{n\pi x}{L}\right) \mathrm{d}x = \frac{2T_0}{L} \times \frac{L}{n\pi} \left[\cos\left(\frac{n\pi x}{L}\right)\right]_0^L \\ &= \frac{2T_0}{n\pi} \left\{(-1)^n - 1\right\} = \begin{cases} \frac{-4T_0}{n\pi} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases} \end{aligned}$$

Hence

$$T(x,t) = \frac{-4T_0}{\pi} \sum_{n \text{ odd}}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi x}{L}\right) \exp\left\{\frac{-n^2\pi^2\alpha^2 t}{L^2}\right\}. \tag{6.16}$$

Notice that $T(x,t)$ tends to 0 as t tends to ∞ at all points of the rod. ■

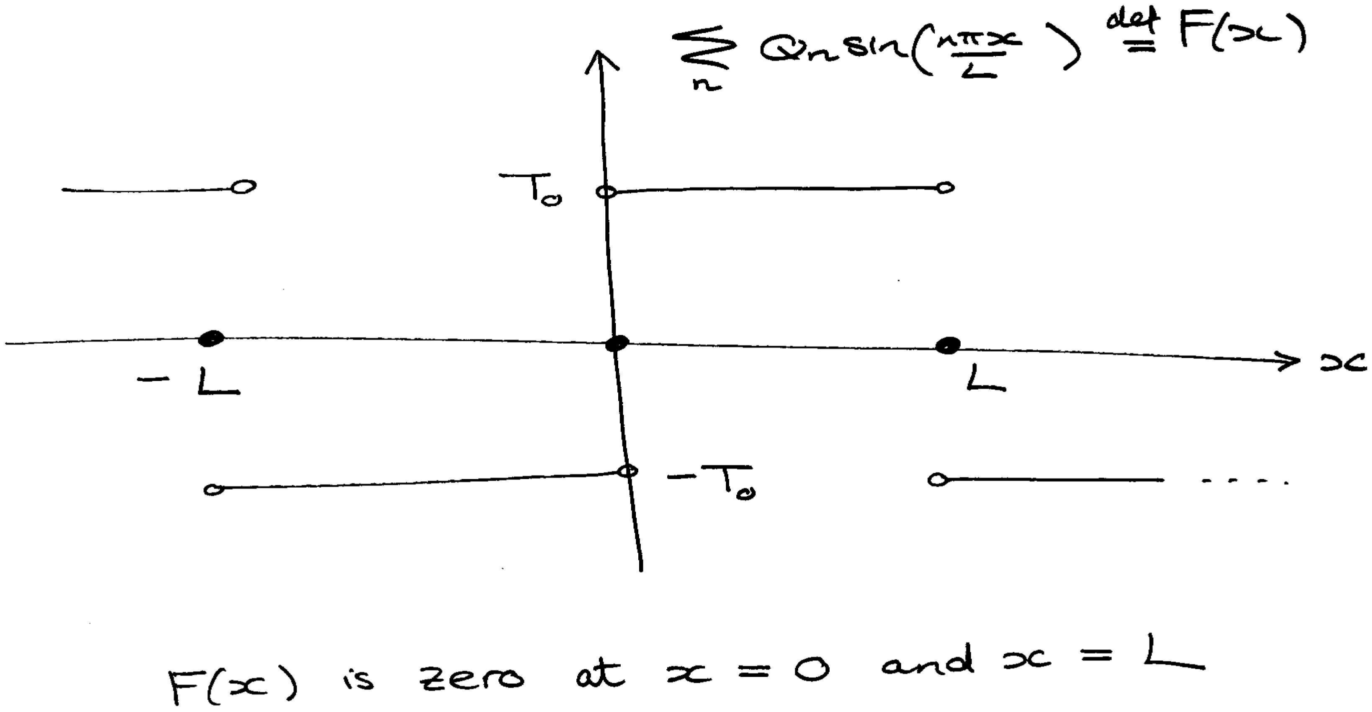


Figure 6-1 The Fourier Series used for T_0 in example solution (6.6) above.

Remark 109 *Note that the Fourier Series*

$$\sum_{n=1}^{\infty} Q_n \sin\left(\frac{n\pi x}{L}\right)$$

cannot be equal to T_0 when $x = 0$ or $x = L$. This Fourier series in fact converges to the function depicted in Figure 6-1. Analogously to example (98) previously, the Fourier series in not continuous. At $x = 0$ it converges to the mean of the limits on approaching $x = 0$ from below and from above, and similarly for other discontinuities. Being VERY careful with the calculation of limits for the expression (6.16) one finds

$$T_0 = \lim_{x\to 0}\lim_{t\to 0} T(x,t) \neq \lim_{t\to 0}\lim_{x\to 0} T(x,t) = 0$$

and such complicated behaviour must be expected given the initial conditions are not consistent with the boundary conditions for the above example.

Example 110 *Suppose now that the temperatures of the ends of the rod are maintained so that*

$$T(0,t) = T_1 \quad \text{and} \quad T(L,t) = T_2.$$

We cannot immediately use separation of variables, because our boundary conditions are no longer homogenous. If we have two solutions which satisfy the heat equation and the above boundary conditions, say T_A and T_B we do NOT have their sum also satisfies the boundary conditions as, for example,

Thus, we cannot find separable solutions and sum them to find a general solution whose summation coefficients can then be fined tune to satisfy the initial conditions.

In direct analogy to differential equations, we can treat this as an inhomogeneous version of the previous problem and just look for a particular solution that satisfies the equation and conditions. One might first consider the steady-state solution that does this as it is easy to find ...

The steady-state heat equation just reads, in this one-dimensional case as

$$T''(x) = 0$$

which has solutions

$$T(x) = Ax + B.$$

Of these the one that meets the boundary conditions is

So if we let

then $U(x,t)$ satisfies the heat equation, the boundary condition

$$U(0,t) = U(L,t) = 0$$

and the initial condition

$$U(x,0) =$$

Utilising Fourier Analysis as in the previous example this problem is now tractable.

6.7 Epilogue

We finish with an aspect of partial derivatives which we have not met previously. The example below is meant as an extremely brief introduction to maximising and minimising functions of many variables which will be gone into in detail later this year.

For the moment let us consider some aspects of one-variable extrema problems. If we are looking to maximise or minimise a differentiable function $f(x)$ on an interval $[a,b]$ then the maxima and minima may arise in one of two ways:

- at **internal** extrema — this is when the function takes its maximum or minimum at a point in (a,b) and in this case $f'(x) = 0$ at such a point;
- at **external** extrema — this is when the function takes its maximum or minimum at either $x = a$ or $x = b$ and it need no longer be the case that $f'(x) = 0$.

By way of example consider the function

$$f(x) = \cos x \quad \text{for } 1 \leqslant x \leqslant 4.$$

It is clear, from the graph below, that the function takes its maximum (in this interval) at $x = \pi/2$ which is an internal maximum, and takes its minimum at $x = 4$ which is an external maximum.

There are various ways to approach the problem below, including purely geometric ones, but we will treat this as an extremum problem of a two variable problem.

Example 111 *Given a triangle with angles A, B, C, none of which is obtuse, show that*

$$2 < \sin A + \sin B + \sin C \leqslant \frac{3\sqrt{3}}{2}.$$

Solution. At first glance this may seem like a three variable problem but of course $A+B+C=\pi$ and so in fact all possibilities are determined by A and B alone. Now (A, B) cannot take all possible values in the plane; rather it must be the case that

$$0 < A < \frac{\pi}{2} \qquad 0 < B < \frac{\pi}{2}$$

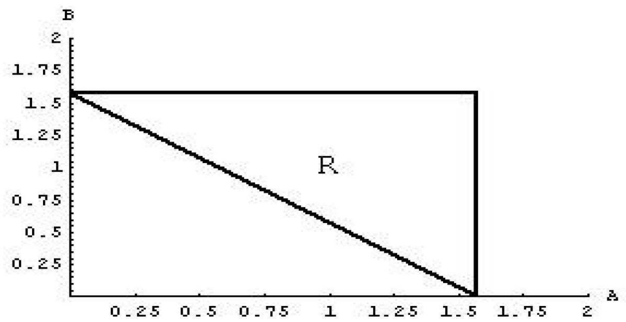
and as $C = \pi - A - B$ is also not obtuse it follows that

$$\frac{\pi}{2} < A + B < \pi.$$

These three inequalities mean that we only need to treat those triangles (A, B) which lie in the interior of

$$R = \left\{(A, B) : 0 \leqslant A, B \leqslant \frac{\pi}{2} \leqslant A + B \leqslant \pi\right\}$$

which is a triangular region of the AB-plane sketched in the diagram below.



Thus we consider the function

$$f(A, B) = \sin A + \sin B + \sin(\pi - A - B) = \sin A + \sin B + \sin(A + B)$$

in the region R.

- **Internal Extrema:** We need to find the critical points (if any) and thus require $\partial f/\partial A = \partial f/\partial B = 0$. Given

$$\begin{aligned}\frac{\partial f}{\partial A} &= \cos A + \cos(A+B) = 0,\\ \frac{\partial f}{\partial B} &= \cos B + \cos(A+B) = 0.\end{aligned}$$

we have

$$\cos A = \cos B = -\cos(A+B) = \cos(\pi - A - B)$$

and thus that $A = B = \pi - A - B$ for the allowed values of A, B.

It is a straightforward matter to calculate f_{AA}, f_{BB}, f_{AB} when $A = B = \pi/3$, and thus to show that this critical point is a maximum. Hence, there is a maximum when the triangle is equilateral and for such triangles $f = 3\sqrt{3}/2$. Given this is the only internal extremum, $3\sqrt{3}/2$ is the maximum value that f can take within the interior of R.

- **External Extrema:** instead of, as in the one-variable case, having two endpoints to check, our boundary now comprises three sides of a triangle. These are given by:

$$\begin{aligned}A &= \pi/2, \qquad 0 \leqslant B \leqslant \pi/2,\\ B &= \pi/2, \qquad 0 \leqslant A \leqslant \pi/2,\\ A+B &= \pi/2, \qquad 0 \leqslant A, B \leqslant \pi/2.\end{aligned}$$

We will treat the first side and set $A = \pi/2$. Then

$$\begin{aligned}f\left(\frac{\pi}{2}, B\right) &= \sin\frac{\pi}{2} + \sin B + \sin\left(\frac{\pi}{2} + B\right)\\ &= 1 + \sin B + \cos B\\ &= 1 + \sqrt{2}\cos\left(B - \frac{\pi}{4}\right).\end{aligned}$$

We can see then that, as B varies over $[0, \pi/2]$ then f is minimal at $B = 0$ and $\pi/2$ at which points $f = 2$ and has a maximum of $1 + \sqrt{2}$ on this side (and this is less than $3\sqrt{3}/2$). Note that when $A = \pi/2$ and $B = 0$ or $\pi/2$ we don't really have a triangle, in the normal sense of the word — rather we have two parallel lines which meet at infinity.

If we now wished we could deal with the remaining two sides similarly, but it is easier to note that they will be the same by the symmetry of the situation — having a right angle at A will result in the same triangles as having a right angle at B or C.

So finally we have our maximum and minimum for f, and noting that the minimum is not actually achieved by a proper triangle we have

$$2 < \sin A + \sin B + \sin C \leqslant \frac{3\sqrt{3}}{2}$$

for any triangle with no obtuse angles.