May 1, 2009

Mathematical Techniques: Revision Notes

Dr A. J. Bevan,

These notes contain the core of the information conveyed in the lectures. They are not a substitute for attending the lectures and none of the examples covered are reproduced here. Worked examples of the techniques described in this note can be found in the tutorial question/solution material provided on the course web site.

1 Rates of change

The change in y with respect to x between these two points on the curve is given by the ratio

$$\begin{aligned} \frac{\delta y}{\delta x} &= \frac{y_2 - y_1}{x_2 - x_1}, \\ &= \frac{y(x+\delta x) - y(x)}{\delta x}. \end{aligned} \tag{1.1}$$

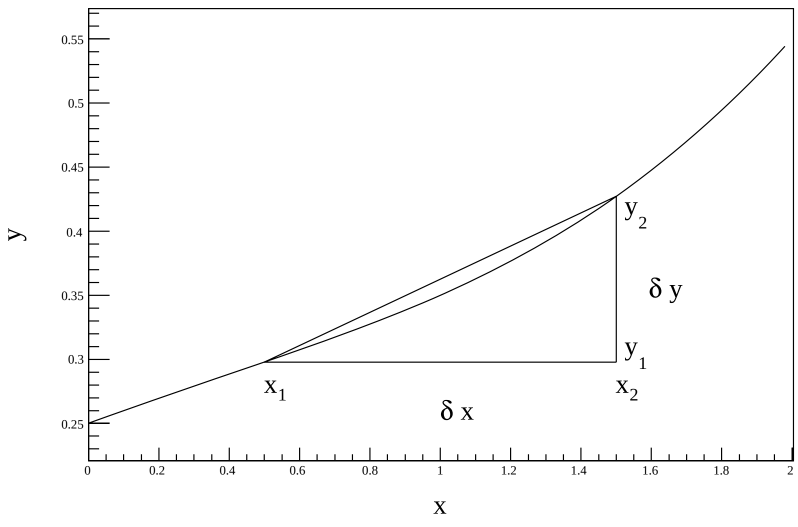

Figure 1: The function $y = f(x)$.

If we take the limit as δx tends to zero we obtain

$$\begin{aligned} \lim_{\delta x \to 0} \left(\frac{\delta y}{\delta x} \right) &= \frac{dy}{dx}, \\ &= \lim_{\delta x \to 0} \left(\frac{y(x+\delta x) - y(x)}{\delta x} \right). \end{aligned} \tag{1.2}$$

which is exact. In general the derivative obtained will be a function of x and can itself be differentiated in order to obtain higher order derivatives.

1.1 Standard Derivatives

It is tedious to work out the derivatives of common functions using Eq. (1.2) each time, and in practice we rely on tables or books of 'standard derivatives'. Table 1 lists a number of useful results.

Table 1: Table of standard derivatives.

$y = f(x)$	$\frac{dy}{dx}$
x^n	nx^{n-1}
e^x	e^x
e^{kx}	ke^{kx}
a^x	$a^x \ln a$
$\ln x$	$\frac{1}{x}$
$\log_a x$	$\frac{1}{x \ln a}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\sec^2 x$
$\cot x$	$-\mathrm{cosec}^2 x$
$\sec x$	$\sec x \tan x$
$\mathrm{cosec} x$	$-\mathrm{cosec} x \cot x$
$\sinh x$	$\cosh x$
$\cosh x$	$\sinh x$
$\tanh x$	$1/\cosh^2 x$
$\sin^{-1} x$	$\frac{1}{\sqrt{1-x^2}}$
$\cos^{-1} x$	$\frac{-1}{\sqrt{1-x^2}}$
$\tan^{-1} x$	$\frac{1}{1+x^2}$
$\sinh^{-1} x$	$\frac{1}{\sqrt{x^2+1}}$
$\cosh^{-1} x$	$\frac{1}{\sqrt{x^2-1}}$
$\tanh^{-1} x$	$\frac{1}{1-x^2}$

2 Chain Rule

Consider the function $y = f(u)$, where $u = g(x)$. Here y is a function of u, which itself is a function of some other variable x. The chain rule states that for such a function

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}. \quad (2.1)$$

3 Differentiation of a Products

The rule for the differentiation of the product of two functions $u(x)$ and $v(x)$ is

$$\frac{d}{dx}(uv) = v\frac{du}{dx} + u\frac{dv}{dx}, \tag{3.1}$$

where it is assumed that both u and v are differentiable with respect to x.

4 Differentiation of Quotients

The rule for the differentiation of the ratio of two functions $u(x)$ and $v(x)$ is

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}, \tag{4.1}$$

where it is assumed that both u and v are differentiable with respect to x.

5 Logarithmic Differentiation

If one has a more complicated function, such as

$$y = \frac{u^n(x)v^m(x)}{h^p(x)}, \tag{5.1}$$

while it is possible to recursively use the quotient and product rules to compute the derivative, it is possible, and often convenient to use natural logarithms to simplify the problem at hand. Taking the log of both sides of Eq. (5.2) one obtains

$$\ln y = n\ln u(x) + m\ln v(x) - p\ln h(x), \tag{5.2}$$

which can be differentiated, noting that y is a function of x so that

$$\frac{d}{dx}(\ln y) = \frac{1}{y}\frac{dy}{dx}, \tag{5.3}$$

thus

$$\frac{1}{y}\frac{dy}{dx} = \frac{n}{u}\frac{du}{dx} + \frac{m}{v}\frac{dv}{dx} - \frac{p}{h}\frac{dh}{dx}, \tag{5.4}$$

which can be used to determine the derivative as a function of x given the form of u, v and h.

6 Implicit Differentiation

It is possible to differentiate functions that are implicitly relating y to x using the chain rule (as with the previous example). For example, consider

$$x^2 + y^2 = R^2, \tag{6.1}$$

which describes a circle. The derivative of this equation is

$$2x + 2y\frac{dy}{dx} = 0, \tag{6.2}$$

so

$$\frac{dy}{dx} = -\frac{x}{y}. \tag{6.3}$$

7 Parametric Functions

If we have a function defined by

$$y = h(\theta) \text{ and } x = g(\theta), \tag{7.1}$$

then we can compute the derivative of y with respect to x by using the chain rule, and noting that

$$\frac{dy}{dx} = \frac{dy}{d\theta} \times \frac{d\theta}{dx}. \tag{7.2}$$

assuming that it is possible do differentiate both h and g with respect to θ.

8 Tangents and normals to a curve

Consider a straight line where

$$y = mx + C, \tag{8.1}$$

and

$$m = \frac{dy}{dx}, \tag{8.2}$$

which is the slope of the line. Given a point on the line (x_1, y_1) it is possible to compute the equation of the line by noting that

$$y - y_1 = m(x - x_1), \tag{8.3}$$
$$y = mx - mx_1 + y_1, \tag{8.4}$$

which is the equation of the tangent to a curve with gradient m at the point (x_1, y_1).

The equation of the normal to a curve at (x_1, y_1) can be obtained by recalling that the gradient of the normal to a curve is $-1/m$. So the equation of the normal to a curve is just

$$y = -\frac{1}{m}x + \frac{1}{m}x_1 + y_1. \tag{8.5}$$

9 Differentiating inverse trigonometric functions

The table of standard derivatives (Table 1) lists the derivatives of the trigonometric functions. These can be used in order to derive the rules for differentiating inverse trigonometric functions. If we consider the equation

$$y = \arcsin(x),$$

we are able to take the sine of both sides to obtain

$$\sin(y) = x.$$

We can now differentiate x with respect to y which yields

$$\begin{aligned}\frac{dx}{dy} &= \cos(y),\\ &= \sqrt{1-x^2},\end{aligned}$$

using $\sin^2(y) + \cos^2(y) = 1$. We can now readily obtain the derivative of $\arcsin(x)$ which is

$$\frac{dy}{dx} = \frac{1}{\sqrt{1-x^2}}$$

It is possible to show that the derivative of $y = \arccos(x)$ is

$$\frac{dy}{dx} = -\frac{1}{\sqrt{1-x^2}}$$

and similar steps can be taken in order to show that for $y = \arctan(x)$

$$\frac{dy}{dx} = \frac{1}{1+x^2}.$$

10 Radius of curvature

A useful property of a curve is the radius of curvature R. The radius of curvature is related to the change in direction $\delta\theta$ of the curve as a function of the change in position δs along the curve. The meaning of this quantity can be seen from the following: In the limit that δs tends to zero, the arc segment of length ds for the function $y = f(x)$ forms part of the circumference of a circle (See Fig. 2). The radius of this circle is R, and the center of the radius of curvature is the point (x_c, y_c). We can relate ds to R via the angle subtended by the arc

$$d\theta = \frac{ds}{R},$$

so,

$$\frac{d\theta}{ds} = R.$$

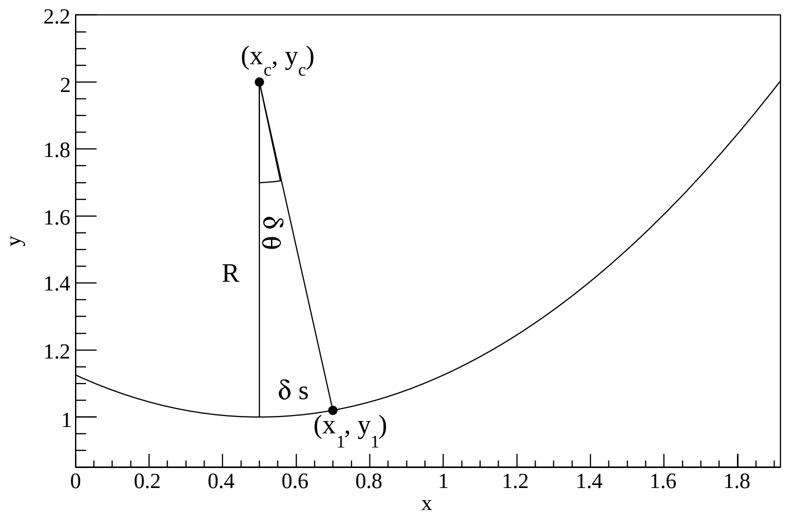

Figure 2: The function $y = f(x)$ showing the radius of curvature R and center of curvature (x_c, y_c) of the function for a point (x_1, y_1) on the curve.

Recall that

$$\frac{dy}{dx} = \tan\theta.$$

If we differentiate both sides of the last equation with respect to the position along the curve s we obtain

$$\begin{aligned}
\frac{d}{ds}\left(\frac{dy}{dx}\right) &= \frac{d}{ds}\left[\tan\theta\right] \\
\frac{d}{dx}\left(\frac{dy}{dx}\right)\frac{dx}{ds} &= \frac{d}{d\theta}\left[\tan\theta\right]\frac{d\theta}{ds} \\
\frac{d^2y}{dx^2}\cos\theta &= \sec^2\theta\frac{d\theta}{ds} \\
\frac{d^2y}{dx^2} &= \sec^3\theta\frac{d\theta}{ds}
\end{aligned}$$

If we recall that $1 + \tan^2\theta = \sec^2\theta$ and that $\tan\theta = y'$, we can substitute for $\sec^3\theta$ to obtain

$$\begin{aligned}
\frac{d^2y}{dx^2} &= (1 + \tan^2\theta)^{3/2}\frac{d\theta}{ds} \\
\frac{d^2y}{dx^2} &= \left[1 + \left(\frac{dy}{dx}\right)^2\right]^{3/2}\frac{d\theta}{ds}.
\end{aligned}$$

We can re-arrange the equation and substitute $\frac{d\theta}{ds}$ for R to obtain

$$R = \frac{\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{3/2}}{\frac{d^2y}{dx^2}} \tag{10.1}$$

On inspection of Eq. (10.1) we can see that the sign of R is given by the sign of the second derivative of the function y. If the second derivative is negative at a point (x, y), then the curve is convex at that point. Conversely if the second derivative is positive at (x, y) then the curve is concave at that point.

The point (x_c, y_c) corresponding to the center of the circle with radius of curvature R is given by

$$x_c = x_1 - R\sin\theta,$$
$$y_c = y_1 + R\cos\theta.$$

11 Stationary points: Maxima, Minima and Inflections

We are able to use the derivatives of a function to obtain more information about features of how the value of a function f changes with x. In particular we can identify the positions where the function f is locally at a maximum or minimum value as illustrated in Figure 3. Generically we call maxima and minima *turning points.* Maxima have a positive gradient of $f(x)$ for $x < x_{maximum}$, and a negative gradient for $x > x_{maximum}$. At the point $x = x_{maximum}$ the gradient is zero. The minima of $f(x)$ can be identified by noting that the gradient is negative for $x < x_{minimum}$, zero for $x = x_{minimum}$ and positive for $x > x_{minimum}$. It is not sufficient to identify minima and maximum solely by $y' = 0$, as we have not considered all of the possibilities as to how the gradient changes. We know that the gradient is changing as we scan through a turning point, and that the sign of the gradient changes sign as we do this. Thus, the value of y'' at a turning point is non-zero. In fact the value of y'' at a turning point is negative for a maximum and positive for a minimum.

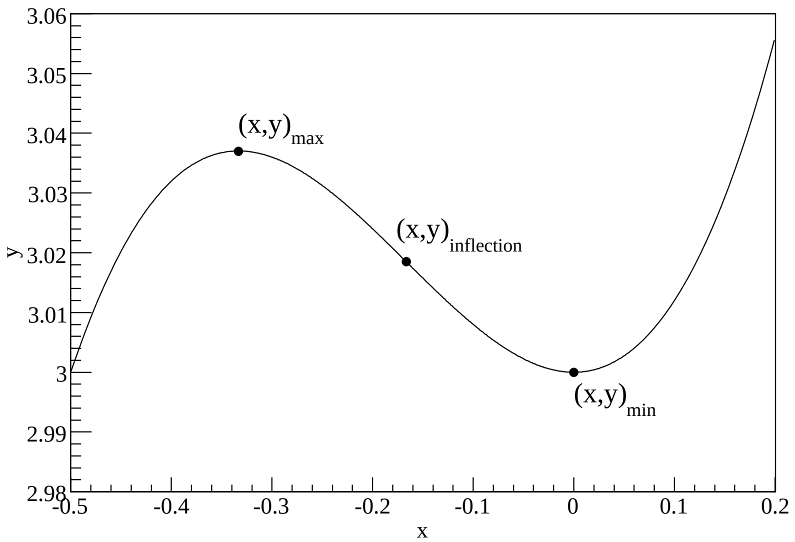

Figure 3: The function $y = f(x)$ illustrating a local maximum, minimum, and point of inflection.

If $y''(x) = 0$ for some value of x, then this point is called a *point of inflection.* In order for the second derivative to be zero at a point of inflection x_I the value of the second derivative has to change sign when we scan from values of $x < x_I$ through $x = x_I$ to $x > x_I$. Collectively turning points and points of inflection are called *stationary points.* The values of $y'(x)$ and $y''(x)$ in the vicinity of stationary points are summarised in Table 2.

Table 2: Derivative values near stationary points: positive values are indicated by + and negative values are indicated by −.

	y'	y''
Maximum		
$x <$ turning point	+	+ or −
x at the turning point	0	−
$x >$ turning point	−	+ or −
Maximum		
$x <$ turning point	−	+ or −
x at the turning point	0	+
$x >$ turning point	+	+ or −
Inflection		
$x <$ turning point	− or +	+ or −
x at the turning point	0	0
$x >$ turning point	+ or −	+ or −

12 Partial Differentiation

Consider the function $z = f(x, y)$, which can be represented as a surface in three dimensions (See Fig. 4). The variables x and y are orthogonal, which is another way of saying that they are independent. This means if we want to differentiate z with respect to x, we can consider y as a constant while doing so. Similarly, if we want to differentiate z with respect to y, we can consider x as a constant. When we do this, we use slightly different notation than when differentiating a function of a single variable. We use the *curly d* symbol '∂' called '*del*' to indicate a partial derivative, i.e. that we are differentiating a function of more than one variable. So we can write

$$\frac{\partial z}{\partial x}$$

as the partial derivative of z with respect to x. This quantity is the rate of change of z with respect to x for a constant y. Similarly

$$\frac{\partial z}{\partial y}$$

is the partial derivative of z with respect to y, or the rate of change of z with respect to y for a constant x.

Higher derivatives of z can be calculated, but it is possible to differentiate either with respect to y or with respect to x, so we have four choices of the second derivative of z. If we differentiate $\partial z/\partial x$ we can obtain

$$\frac{\partial^2 z}{\partial x^2},$$
$$\frac{\partial^2 z}{\partial y \partial x},$$

by differentiating with respect to x or y, respectively. Similarly we obtain

$$\frac{\partial^2 z}{\partial y^2},$$
$$\frac{\partial^2 z}{\partial x \partial y},$$

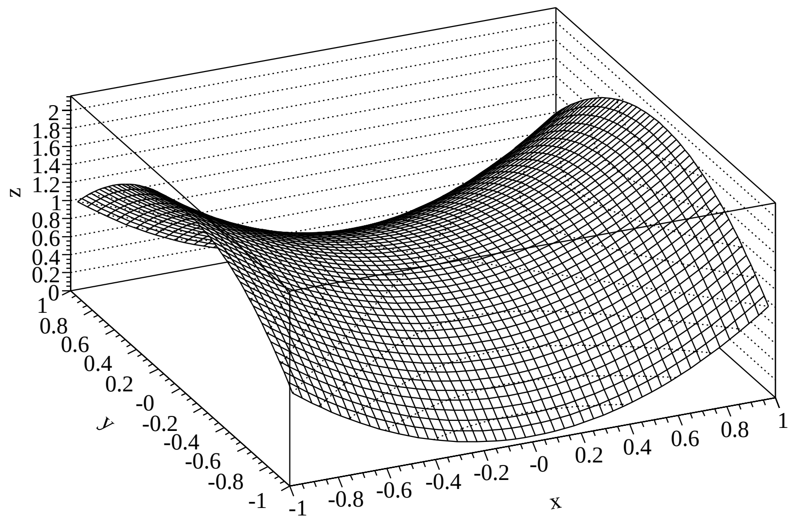

Figure 4: The curve $f(x, y)$ about a saddle point.

by differentiating $\partial z/\partial y$ with respect to y or x, respectively. A consequence of the orthogonality of y and x is that the same result is obtained when differentiating z by y then x, and when differentiating z by x then y, to put this another way:

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

13 Revisiting stationary points

We can extend the discussion in Section 11 to functions of more than one variable. For a function $f(x, y)$ it is possible to use the tests summarised in Table 2 to identify stationary point either in x or in y. In order to identify local maximum we require that a function $f(x, y)$ has a maximum in both x and y. Similarly for a point to be at a local minimum of the function we require that both x and y are at minimum. As shown in Figure 4 it is possible that a function has a minimum in x and a maximum in y (or visa versa) at a given point (x_s, y_s). When this situation arises we call the point (x_s, y_s) a saddle point.

14 Total Derivative

For some function $z = f(x, y)$, it can be useful to know how z changes for small changes in both x and y. If we recall that x and y are independent variables, any small change δx in the variable x is independent of a change δy in the variable y. Let us consider the case where y is constant, and x changes by δx. We can write the change in z simply as

$$\delta z = \frac{\partial z}{\partial x} \delta x.$$

If we now consider x to be constant, and allow y to change by some small amount δy, we are able to write down the change in z as

$$\delta z = \frac{\partial z}{\partial y}\delta y.$$

Each of these solutions results from the fact that x and y are independent, and $z = f(x, y)$ can be reduced into a one-dimensional problem. Now we are ready to consider the general case where both x and y independently change by some small amount, and calculate the change in z which is

$$\delta z = \frac{\partial z}{\partial x}\delta x + \frac{\partial z}{\partial y}\delta y. \tag{14.1}$$

This is called the total derivative, and it is the sum of the change in z expected as a result of a small change in x, and the change in z expected for a small change in y.

14.0.1 Rates of Change with Respect to Time

Consider the change z when x and y change in an interval of time δt; Eq. (14.1) becomes (dividing by δt)

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

In the limit where $\delta t \to 0$:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}. \tag{14.2}$$

15 Change of Variables

So in general for some function $x = f(x, y)$, where $x = x(u)$, and $y = y(u)$

$$\frac{dz}{du} = \frac{\partial z}{\partial x}\frac{dx}{du} + \frac{\partial z}{\partial y}\frac{dy}{du},$$

We can go a step further. If some function is given by $z = f(x, y)$, where $x = h(u, v)$, and $y = g(u, v)$, where u and v are two orthogonal variables, then z is also a function of u and v and we can write:

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}, \tag{15.1}$$

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}. \tag{15.2}$$

NOTE: With the function $y = f(x)$ it is possible to invert a derivative as the relation

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}.$$

It can be convenient to use this relation if it is simpler to invert and differentiate a function with respect to y than with respect to x. The same is not true with partial differentiation! We can see this by considering derivatives of the following function $z = x^2 + y$. The derivative of z with respect to x is simply $2x$. However in order to differentiate x with respect to z, we must first re-write the function as $x = \sqrt{z - y}$. The derivative of x with respect to z is $1/(2\sqrt{z - y})$.

15.1 Changing between polar and cartesian co-ordinates

We are able to write a function $z = f(x, y)$ in terms of polar coordinates in terms of the distance from the origin r, and an angle subtended between the radial distance to a point on the function, the origin and the $x-$axis θ. The angle θ has a value between zero and 2π. We can determine expressions for the cartesian coordinates x and y in terms of r and θ. These are simply

$$\begin{aligned} x &= r\cos(\theta), \\ y &= r\sin(\theta). \end{aligned}$$

We can write the derivative of z in terms of r and θ by noting that

$$\begin{aligned} \frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}, \\ \frac{\partial f}{\partial \theta} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta}. \end{aligned}$$

Where,

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

and,

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

Using these results we can write

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

We are able to use this result to determine the derivative of f with respect to r and θ for any differentiable function $f = f(x, y)$.