Appendix A

Background material on calculus (B)

A.1 Notation and preliminary remarks

We start with a discussion about notation and some of the basic ideas used throughout this module.

A real function of a single real variable, $f$, is a rule that maps a real number $x$ to a single real number $y$. This operation can be denoted in a variety of ways. The approach of scientists is to write $y = f(x)$ or just $y(x)$, and the symbols $y$, $y(x)$, $f$ and $f(x)$ are all used to represent the function. Mathematics uses the more formal and precise notation $f : X \rightarrow Y$, where $X$ and $Y$ are subsets of the real line: the set $X$ is named the domain, or the domain of definition of $f$, and set $Y$ the codomain. With this notation the symbol $f$ denotes the function and the symbol $f(x)$ the value of the function at the point $x$. In applications this distinction is not always made and both $f$ and $f(x)$ are used to denote the function. In this text we shall frequently use the Leibniz notation, $f(x)$, and its extensions, because it generally provides a clearer picture and is helpful for algebraic manipulations, such as when changing variables and integrating by parts.

Moreover, in the sciences the domain and codomain are frequently omitted, either because they are 'obvious' or because they are not known. But, perversely, the scientist, by writing $y = f(x)$, often distinguishes between the two variables $x$ and $y$, by saying that $x$ is the independent variable and that $y$ is the dependent variable because it depends upon $x$. This labelling can be confusing, because the role of variables can change, but it is also helpful because in physical problems different variables can play quite different roles: for instance, time is normally an independent variable.

In pure mathematics the term graph is used in a slightly specialised manner. A graph is the set of points $(x, f(x))$: this is normally depicted as a line in a plane using rectangular Cartesian coordinates. In other disciplines the whole figure is called the graph, not the set of points, and the graph may be a less restricted shape than those defined by functions; an example is shown in figure A.5 (page 364).

Almost all the ideas associated with real functions of one variable generalise to functions of several real variables, but notation needs to be developed to cope with this
extension. Points in $\mathbb{R}^n$ are represented by $n$-tuples of real numbers $(x_1, x_2, \ldots, x_n)$. It is convenient to use bold faced symbols, $\mathbf{x}$, and so on, to denote these points, so $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and we shall write $\mathbf{x}$ and $(x_1, x_2, \ldots, x_n)$ interchangeably. In handwritten text a bold character, $\mathbf{x}$, is usually denoted by an underline, _x_.

A function $f(x_1, x_2, \ldots, x_n)$ of $n$ real variables, defined on $\mathbb{R}^n$, is a map from $\mathbb{R}^n$, or a subset, to $\mathbb{R}$, written as $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Where we use bold face symbols like $f$ or $\phi$ to refer to functions, it means that the image under the function $f(x)$ or $\phi(y)$ may be considered as vector in $\mathbb{R}^m$ with $m \geq 2$, so $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$; in this module normally $m = 1$ or $m = n$. Although the case $m = 1$ will not be excluded when we use a bold face symbol, we shall continue to write $f$ and $\phi$ where the functions are known to be real valued and not vector valued. We shall also write without further comment $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$, so that the $f_i$ are the $m$ component functions, $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, of $f$.

On the real line the distance between two points $x$ and $y$ is naturally defined by $|x - y|$. A point $x$ is in the open interval $(a, b)$ if $a < x < b$, and is in the closed interval $[a, b]$ if $a \leq x \leq b$. By convention, the intervals $(-\infty, a)$, $(b, \infty)$, and $(-\infty, \infty) = \mathbb{R}$ are also open intervals. Here, $(-\infty, a)$ means the set of all real numbers strictly less than $a$. The symbol $\infty$ for ‘infinity’ is not a number, and its use here is conventional. In the language and notation of set theory, we can write $(-\infty, a) = \{x \in \mathbb{R} : x < a\}$, with similar definitions for the other two types of open interval. One reason for considering open sets is that the natural domain of definition of some important functions is an open set. For example, the function $\ln x$ as a function of one real variable is defined for $x \in (0, \infty)$.

The space of points $\mathbb{R}^n$ is an example of a linear space. Here the term linear has the normal meaning that for every $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^n$, and for every real $\alpha$, $\mathbf{x} + \mathbf{y}$ and $\alpha \mathbf{x}$ are in $\mathbb{R}^n$. Explicitly,

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

and

$$\alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n).$$

Functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ may also be added and multiplied by real numbers. Therefore a function of this type may be regarded as a vector in the vector space of functions — though this space is not finite dimensional like $\mathbb{R}^n$.

In the space $\mathbb{R}^n$ the distance $|\mathbf{x}|$ of a point $\mathbf{x}$ from the origin is defined by the natural generalisation of Pythagoras’ theorem, $|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$. The distance between two vectors $\mathbf{x}$ and $\mathbf{y}$ is then defined by

$$|\mathbf{x} - \mathbf{y}| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}. \quad (A.1)$$

This is a direct generalisation of the distance along a line, to which it collapses when $n = 1$.

This distance has the three basic properties

(a) $|\mathbf{x}| \geq 0$ and $|\mathbf{x}| = 0$ if and only if $\mathbf{x} = 0$,

(b) $|\mathbf{x} - \mathbf{y}| = |\mathbf{y} - \mathbf{x}|$,

(c) $|\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}| \geq |\mathbf{x} - \mathbf{z}|$, (Triangle inequality) .

In the more abstract spaces, such as the function spaces we need later, a similar concept of a distance between elements is needed. This is named the norm and is a map from
two elements of the space to the positive real numbers and which satisfies the above three rules. In function spaces there is no natural choice of the distance function and we shall see in chapter 4 that this flexibility can be important.

For functions of several variables, that is, for functions defined on sets of points in \( R^n \), the direct generalization of open interval is an open ball.

**Definition A.1** The open ball \( B_r(a) \) of radius \( r \) and centre \( a \in R^n \) is the set of points

\[
B_r(a) = \{ x \in R^n : |x - a| < r \},
\]

Thus the ball of radius 1 and centre \((0,0)\) in \( R^2 \) is the interior of the unit circle, not including the points on the circle itself. And in \( R \), the ‘ball’ of radius 1 and centre 0 is the open interval \((-1,1)\). However, for \( R^2 \) and for \( R^n \) for \( n > 2 \), open balls are not quite general enough. For example, the open square

\[
\{(x, y) \in R^2 : |x| < 1, \ |y| < 1 \}
\]

is not a ball, but in many ways is similar. (You may know for example that it may be mapped continuously to an open ball.) It turns out that the most convenient concept is that of open set which we can now define.

**Definition A.2** Open sets. A set \( U \) in \( R^n \) is said to be open if for every \( x \in U \) there is an open ball \( B_r(a) \), \( r > 0 \), wholly contained within \( U \) and which contains \( x \).

In other words, every point in an open set lies in an open ball contained in the set. Any open ball is in many ways like the whole of the space \( R^n \) — it has no isolated or missing points. Also, every open set is a union of open balls (obviously). Open sets are very convenient and important in the theory of functions, but we cannot study the reasons here. A full treatment of open sets can be found in books on topology.\(^1\) Open balls are not the only type of open sets and it is not hard to show that the open square, \( \{(x, y) \in R^2 : |x| < 1, |y| < 1 \} \), is in fact an open set, according to the definition we gave; and in a similar way it can be shown that the set \( \{(x, y) \in R^2 : (x/a)^2 + (y/b)^2 < 1 \} \), which is the interior of an ellipse, is an open set.

**Exercise A.1 (B)** Show that the open square is an open set by constructing explicitly for each \((x, y)\) in the open square \( \{(x, y) \in R^2 : |x| < 1, |y| < 1 \} \) a ball containing \((x, y)\) and lying in the square.

---

A.2 Functions of a real variable

A.2.1 Introduction

In this section we introduce important ideas pertaining to real functions of a single real variable, although some mention is made of functions of many variables. Most of the ideas discussed should be familiar from earlier courses in elementary real analysis or calculus, so our discussion is brief and all exercises are optional.

\(^1\) See for example W A Sutherland, *Introduction to Metric and Topological Spaces*, Oxford University Press.
The study of Real Analysis normally starts with a discussion of the real number system and its properties. Here we assume all necessary properties of this number system and refer the reader to any basic text if further details are required: adequate discussion may be found in the early chapters of the texts by Whittaker and Watson\textsuperscript{2}, Rudin\textsuperscript{3} and by Kolmogorov and Fomin\textsuperscript{4}.

\subsection*{A.2.2 Continuity and Limits}

A \textit{continuous} function is one whose graph has no vertical breaks: otherwise, it is discontinuous. The function $f_1(x)$, depicted by the solid line in figure A.1 is continuous for $x_1 < x < x_2$. The function $f_2(x)$, depicted by the dashed line, is discontinuous at $x = c$.

![Figure A.1: Figure showing examples of a continuous function, $f_1(x)$, and a discontinuous function $f_2(x)$.](image)

A function $f(x)$ is continuous at a point $x = a$ if $f(a)$ exists and if, given any arbitrarily small positive number, $\epsilon$, we can find a neighbourhood of $x = a$ such that in it $|f(x) - f(a)| < \epsilon$. We can express this in terms of limits and since a point $a$ on the real line can be approached only from the left or the right a function is continuous at a point $x = a$ if it approaches the same value, independent of the direction. Formally we have

\begin{definition}
\textbf{Continuity}: a function, $f$, is continuous at $x = a$ if $f(a)$ is defined and
\[ \lim_{x \to a} f(x) = f(a). \]
\end{definition}

For a function of one variable, this is equivalent to saying that $f(x)$ is continuous at $x = a$ if $f(a)$ is defined and the left and right-hand limits
\[ \lim_{x \to a^-} f(x) \quad \text{and} \quad \lim_{x \to a^+} f(x), \]
exist \textit{and} are equal to $f(a)$.

If the left and right-hand limits exist but are not equal the function is discontinuous at $x = a$ and is said to have a \textit{simple} discontinuity at $x = a$.

If they both exist and are equal, but do not equal $f(a)$, then the function is said to have a \textit{removable} discontinuity at $x = a$. \hfill \Box

\begin{footnotesize}
\textsuperscript{2}Course of Modern Analysis by E T Whittaker and G N Watson, Cambridge University Press.
\textsuperscript{3}Principles of Mathematical Analysis by W Rudin, McGraw-Hill.
\textsuperscript{4}Introductory Real Analysis by A N Kolmogorov and S V Fomin, Dover.
\end{footnotesize}
A.2. FUNCTIONS OF A REAL VARIABLE

Quite elementary functions exist for which neither limit exists: these are also discontinuous, and said to have a discontinuity of the second kind at \( x = a \), see Rudin (1976, page 94). An example of a function with such a discontinuity at \( x = 0 \) is

\[
f(x) = \begin{cases} 
\sin(1/x) & x \neq 0 \\
0 & x = 0.
\end{cases}
\]

We shall have no need to consider this type of discontinuity in this module, but simple discontinuities will arise.

A function that behaves as

\[
|f(x + \epsilon) - f(x)| = O(\epsilon) \quad \text{as} \quad \epsilon \to 0
\]

is continuous, though the converse is not true, a counterexample being \( f(x) = \sqrt{|x|} \) at \( x = 0 \).

Most functions that occur in the sciences are either continuous or piecewise continuous, which means that the function is continuous except at a discrete set of points. The Heaviside function and the related sgn functions are examples of commonly occurring piecewise continuous functions that are discontinuous. They are defined by

\[
H(x) = \begin{cases} 
1, & x > 0, \\
0, & x < 0,
\end{cases}
\quad \text{and} \quad 
\text{sgn}(x) = \begin{cases} 
1, & x < 0, \\
-1, & x > 0,
\end{cases}
\quad \text{sgn}(x) = -1 + 2H(x). \quad (A.3)
\]

These functions are discontinuous at \( x = 0 \), where they are not normally defined. In some texts these functions are defined at \( x = 0 \); for instance \( H(0) \) may be defined to have the value 0, 1/2 or 1.

If \( \lim_{x \to c} f(x) = A \) and \( \lim_{x \to c} g(x) = B \), then it can be shown that the following (obvious) rules are adhered to:

(a) \( \lim_{x \to c} (\alpha f(x) + \beta g(x)) = \alpha A + \beta B \);

(b) \( \lim_{x \to c} (f(x)g(x)) = AB \);

(c) \( \lim_{x \to c} \frac{f(x)}{g(x)} = \frac{A}{B} \) if \( B \neq 0 \);

(d) if \( \lim_{x \to c} f(x) = f_B \) then \( \lim_{x \to c} f(g(x)) = f_B \).

The value of a limit is normally found by a combination of suitable re-arrangements and expansions. An example of an expansion is

\[
\lim_{x \to 0} \frac{\sinh ax}{x} = \lim_{x \to 0} \frac{ax + \frac{1}{4!}(ax)^3 + O(x^5)}{x} = \lim_{x \to 0} (a + O(x^2)) = a.
\]

An example of a re-arrangement, using the above rules, is

\[
\lim_{x \to 0} \frac{\sinh ax}{\sinh bx} = \lim_{x \to 0} \frac{\sinh ax}{x} \frac{x}{\sinh bx} = \lim_{x \to 0} \frac{\sinh ax}{x} \lim_{x \to 0} \frac{x}{\sinh bx} = \frac{a}{b} \quad (b \neq 0).
\]

Finally, we note that a function that is continuous on a closed interval is bounded above and below and attains its bounds. It is important that the interval is closed; for instance the function \( f(x) = x \) defined in the open interval \( 0 < x < 1 \) is bounded above and below, but does not attain it bounds. This example may seem trivial, but similar difficulties exist in the Calculus of Variations and are less easy to recognise.
Exercise A.2 (B) A function that is finite and continuous for all $x$ is defined by
\[
f(x) = \begin{cases} 
\frac{A}{x} + x + B, & 0 \leq x \leq a, \quad a > 0, \\
\frac{C}{x} + Dx, & a \leq x,
\end{cases}
\]
where $A, B, C, D$ and $a$ are real numbers: if $f(0) = 1$ and $\lim_{x \to \infty} f(x) = 0$, find these numbers.

Exercise A.3 (B) Find the limits of the following functions as $x \to 0$ and $w \to \infty$.
(a) $\frac{\sin ax}{x}$, (b) $\frac{\tan ax}{x}$, (c) $\frac{\sin ax}{\sin bx}$, (d) $\frac{3x + 4}{4x + 2}$, (e) $(1 + \frac{z}{w})^w$.

For functions of two or more variables, the definition of continuity is essentially the same as for a function of one variable. A function $f(x)$ is continuous at $x = a$ if $f(a)$ is defined and
\[
\lim_{x \to a} f(x) = f(a). \tag{A.4}
\]
Alternatively, given any $\epsilon > 0$ there is a $\delta > 0$ such that whenever $|x - a| < \delta$, $|f(x) - f(a)| < \epsilon$.

It should be noted that if $f(x, y)$ is continuous in each variable, it is not necessarily continuous in both variables. For instance, consider the function
\[
f(x, y) = \begin{cases} 
\frac{(x + y)^2}{x^2 + y^2}, & x^2 + y^2 \neq 0, \\
1, & x = y = 0,
\end{cases}
\]
and for fixed $y = \beta \neq 0$ the related function of $x$,
\[
f(x, \beta) = \frac{(x + \beta)^2}{x^2 + \beta^2} = 1 + O(x) \quad \text{as} \quad x \to 0
\]
and $f(x, 0) = 1$ for all $x$: for any $\beta$ this function is a continuous function of $x$. On the line $x + y = 0$, however, $f = 0$ except at the origin so $f(x, y)$ is not continuous along this line. More generally, by putting $x = r \cos \theta$ and $y = r \sin \theta$, $-\pi < \theta \leq \pi$, $r \neq 0$, we can approach the origin from any angle. In this representation $f = 2 \sin^2(\theta + \frac{\pi}{4})$ so on any circle round the origin $f$ takes any value between 0 and 2. Therefore $f(x, y)$ is not a continuous function of both $x$ and $y$.

Exercise A.4 (B) Determine whether or not the following functions are continuous at the origin.
(a) $f = \frac{2xy}{x^2 + y^2}$, (b) $f = \frac{x^2 + y^2}{x^2 - y^2}$, (c) $f = \frac{2x^2y}{x^2 + y^2}$.
Hint: use polar coordinates $x = r \cos \theta$, $y = r \sin \theta$ and consider the limit $r \to 0$.

A.2.3 Monotonic functions and inverse functions

A function is said to be monotonic on an interval if it is always increasing or always decreasing. Simple examples are $f(x) = x$ and $f(x) = \exp(-x)$ which are monotonic.
The gradient of the chord where \( h \neq 0 \), and is given by the formula
\[
\tan \phi = \frac{f(a + h) - f(a)}{h},
\]
where \( h \) is the change of a function, so it is important in the discussion of any geometric idea of the tangent to a curve and to the related concept. The recognition of the intervals on which a given function is strictly monotonic is sometimes important because on these intervals the inverse function exists. For instance the function \( y = e^x \) is monotonic increasing on the whole real line, \( R \), and its inverse is the well known natural logarithm, \( x = \ln y \), with \( y \) on the positive real line.

In general if \( f(x) \) is continuous and strictly monotonic on \( a \leq x \leq b \) and \( y = f(x) \) the inverse function, \( x = f^{-1}(y) \), is continuous for \( f(a) \leq y \leq f(b) \) and satisfies \( y = f(f^{-1}(y)) \). Moreover, if \( f(x) \) is strictly increasing so is \( f^{-1}(y) \).

Complications occur when a function is increasing and decreasing on neighbouring intervals, for then the inverse may have two or more values. For example the function \( f(x) = x^2 \) is monotonic increasing for \( x > 0 \) and monotonic decreasing for \( x < 0 \); hence the relation \( y = x^2 \) has the two familiar inverses \( x = \pm \sqrt{y} \), \( y \geq 0 \). These two inverses are often referred to as the different branches of the inverse; this idea is important because most functions are monotonic only on part of their domain of definition.

**Exercise A.5 (B)**

(a) Show that \( y = 3a^2 x - x^3 \) is strictly increasing for \(-a < x < a\) and that on this interval \( y \) increases from \(-2a^3\) to \(2a^3\).

(b) By putting \( x = 2a \sin \phi \) and using the identity \( \sin^3 \phi = (3 \sin \phi - \sin 3\phi)/4 \), show that the equation becomes
\[
y = 2a^3 \sin 3\phi \quad \text{and hence that} \quad x(y) = 2a \sin \left( \frac{1}{3} \sin^{-1} \left( \frac{y}{2a^3} \right) \right).
\]

(c) Find the inverse for \( x > 2a \). Hint: put \( x = 2a \cosh \phi \) and use the relation \( \cosh^3 \phi = (\cosh 3\phi + 3 \cosh \phi)/4 \).

### A.2.4 The derivative

The notion of the derivative of a continuous function, \( f(x) \), is closely related to the geometric idea of the tangent to a curve and to the related concept of the rate of change of a function, so is important in the discussion of anything that changes. This geometric idea is illustrated in figure A.2: here \( P \) is a point with coordinates \((a, f(a))\) on the graph and \( Q \) is another point on the graph with coordinates \((a + h, f(a + h))\), where \( h \) may be positive or negative.

The gradient of the chord \( PQ \) is \( \tan \phi \) where \( \phi \) is the angle between \( PQ \) and the \( x \)-axis, and is given by the formula
\[
\tan \phi = \frac{f(a + h) - f(a)}{h},
\]
Figure A.2: Illustration showing the chord PQ and the tangent line at P.

If the graph in the vicinity of $x = a$ is represented by a smooth line, then it is intuitively obvious that the chord $PQ$ becomes closer to the tangent at $P$ as $h \to 0$; and in the limit $h = 0$ the chord becomes the tangent. Hence the gradient of the tangent is given by the limit

$$
\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.
$$

This limit, provided it exists, is named the derivative of $f(x)$ at $x = a$ and is commonly denoted either by $f'(a)$ or $\frac{df}{dx}$. Thus we have the formal definition:

**Definition A.5 The derivative:** A function $f(x)$, defined on an open interval $U$ of the real line, is differentiable at $x \in U$ and has the derivative $f'(x)$ if

$$
f'(x) = \frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}, \quad (A.5)
$$

exists.

If the derivative exists at every point in the open interval $U$ the function $f(x)$ is said to be differentiable in $U$: in this case it may be proved that $f(x)$ is also continuous. However, a function that is continuous at $a$ need not be differentiable at $a$: indeed, it is possible to construct functions that are continuous everywhere but differentiable nowhere; such functions are encountered in the mathematical description of Brownian motion.

Combining the definition of $f'(x)$ and the definition of the $\theta$ order notation on page 43 shows that a differentiable function satisfies

$$
f(x + h) = f(x) + hf'(x) + o(h). \quad (A.6)
$$

The formal definition, equation (A.5), of the derivative can be used to derive all its useful properties, but the physical interpretation, illustrated in figure A.2, provides a more useful way to generalise it to functions of several variables.

The tangent line to the graph $y = f(x)$ at the point $a$, which we shall consider to be fixed for the moment, has slope $f'(a)$ and passes through $f(a)$. These two facts determine the derivative completely. The equation of the tangent line can be written in parametric form as $p(h) = f(a) + f'(a)h$. Conversely, given a point $a$, and the equation of the tangent line at that point, the derivative, in the classical sense of the definition A.5, is simply the slope, $f'(a)$, of this line. So the information that the
derivative of \( f \) at \( a \) is \( f'(a) \) is equivalent to the information that the tangent line at \( a \) has equation \( p(h) = f(a) + f'(a)h \). Although the classical derivative, equation (A.5), is usually taken to be the fundamental concept, the equivalent concept of the tangent line at a point could be considered equally fundamental — perhaps more so, since a tangent is a more intuitive idea than the numerical value of its slope. This is the key to successfully defining the derivative of functions of more than one variable.

From the definition A.5 the following useful results follow. If \( f(x) \) and \( g(x) \) are differentiable on the same open interval and \( \alpha \) and \( \beta \) are constants then

\[
\begin{align*}
(a) & \quad \frac{d}{dx}(\alpha f(x) + \beta g(x)) = \alpha f'(x) + \beta g'(x), \\
(b) & \quad \frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x), \quad \text{(The product rule)} \\
(c) & \quad \frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}, \quad g(x) \neq 0. \quad \text{(The quotient rule)}
\end{align*}
\]

We leave the proof of these results to the reader, but note that the differential of \( 1/g(x) \) follows almost trivially from the definition A.5 and exercise A.10, so that the third expression is a simple consequence of the second.

The other important result is the chain rule concerning the derivative of composite functions. Suppose that \( f(x) \) and \( g(x) \) are two differentiable functions and a third is formed by the composition,

\[ F(x) = f(g(x)), \quad \text{sometimes written as} \quad F = f \circ g, \]

which we assume to exist. Then the derivative of \( F(x) \) can be shown, as in exercise A.14, to be given by

\[
\frac{dF}{dx} = \frac{df}{dg} \times \frac{dg}{dx} \quad \text{or} \quad F'(x) = f'(g(x))g'(x). \tag{A.7}
\]

This formula is named the chain rule. Note how the prime-notation is used: it denotes the derivative of the function with respect to the argument shown, not necessarily the original independent variable, \( x \). Thus \( f'(g(x)) \) is not the derivative of \( F(x) \); it means the derivative \( f'(x) \) with \( x \) replaced by \( g(x) \).

A simple example should make this clear: suppose \( f(x) = \sin x \) and \( g(x) = 1/x, x > 0 \), so \( F(x) = \sin(1/x) \). The chain rule gives

\[
\frac{dF}{dx} = \frac{d}{dg}(\sin g) \times \frac{d}{dx}\left(\frac{1}{x}\right) = \cos g \times \left(-\frac{1}{x^2}\right) = -\frac{1}{x^2} \cos \left(\frac{1}{x}\right).
\]

A.6 (B) Find the derivative of the following functions

\[
\begin{align*}
(a) & \quad \sqrt{(a-x)(b+x)}, \quad (b) \quad \sqrt{a\sin^2 x + b\cos^2 x}, \quad (c) \quad \cos(x^3) \cos x, \quad (d) \quad x^a.
\end{align*}
\]
Exercise A.7 (B) If \( y = \sin x \) for \(-\pi/2 \leq x \leq \pi/2\) show that \( \frac{dx}{dy} = \frac{1}{\sqrt{1 - y^2}} \).

Exercise A.8 (B) (a) If \( y = f(x) \) has the inverse \( x = g(y) \), show that \( f'(x)g'(y) = 1 \), that is
\[
\frac{dx}{dy} = \left( \frac{dy}{dx} \right)^{-1}.
\]
(b) Express \( \frac{d^2x}{dy^2} \) in terms of \( \frac{dy}{dx} \) and \( \frac{d^2y}{dx^2} \).

Clearly, if \( f'(x) \) is differentiable, it may be differentiated to obtain the second derivative, which is denoted by
\[
f''(x) \quad \text{or} \quad \frac{d^2f}{dx^2}.
\]
This process can be continued to obtain the functions
\[
f, \quad \frac{df}{dx}, \quad \frac{d^2f}{dx^2}, \quad \frac{d^3f}{dx^3}, \ldots, \quad \frac{d^{n-1}f}{dx^{n-1}}, \quad \frac{d^n f}{dx^n}, \ldots
\]
where each member of the sequence is the derivative of the preceding member,
\[
\frac{d^p f}{dx^p} = \frac{d}{dx} \left( \frac{d^{p-1} f}{dx^{p-1}} \right), \quad p = 2, 3, \ldots
\]
The prime notation becomes rather clumsy after the second or third derivative, so the most common alternative is
\[
\frac{d^p f}{dx^p} = f^{(p)}(x), \quad p \geq 2,
\]
with the conventions \( f^{(1)}(x) = f'(x) \) and \( f^{(0)}(x) = f(x) \). Care is needed to distinguish between the \( p \)th derivative, \( f^{(p)}(x) \), and the \( p \)th power, denoted by \( f(x)^p \) and sometimes \( f^p(x) \) — the latter notation should be avoided if there is any danger of confusion.

Functions for which the \( n \)th derivative is continuous are said to be \( n \)-differentiable and to belong to class \( C_n \) : the notation \( C_n(U) \) means the first \( n \) derivatives are continuous on the interval \( U \): the notation \( C_n(a, b) \) or \( C_n[a, b] \), with obvious meaning, may also be used. The term smooth function describes functions belonging to \( C_\infty \), that is functions, such as \( \sin x \), having all derivatives; we shall, however, use the term sufficiently smooth for functions that are sufficiently differentiable for all subsequent analysis to work, when more detail is deemed unimportant.

In the following exercises some important, but standard, results are derived.

Exercise A.9 (B) If \( f(x) \) is an even (odd) function, show that \( f'(x) \) is an odd (even) function.

Exercise A.10 (B) Show, from first principles using the limit (A.5), that
\[
\frac{d}{dx} \left( \frac{1}{f(x)} \right) = \frac{f'(x)}{f(x)^2},
\]
and that the product rule is true.
Exercise A.11 (B) Leibniz’s rule

If \( h(x) = f(x)g(x) \) show that
\[
\begin{align*}
    h''(x) &= f''(x)g(x) + 2f'(x)g'(x) + f(x)g''(x), \\
    h^{(3)}(x) &= f^{(3)}(x)g(x) + 3f''(x)g'(x) + 3f'(x)g''(x) + f(x)g^{(3)}(x),
\end{align*}
\]
and use induction to derive Leibniz’s rule
\[
h^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)}(x)g^{(k)}(x),
\]
where the binomial coefficients are given by
\[
\binom{n}{k} = \frac{n!}{k!(n-k)!}.
\]

Exercise A.12 (B) Show that
\[
\frac{d}{dx} \ln(f(x)) = \frac{f'(x)}{f(x)}
\]
and hence that if
\[
p(x) = f_1(x)f_2(x) \cdots f_n(x)
\]
then
\[
\frac{p'}{p} = \frac{f'_1}{f_1} + \frac{f'_2}{f_2} + \cdots + \frac{f'_n}{f_n},
\]
provided \( p(x) \neq 0 \). Note that this gives an easier method of differentiating products of three or more factors than repeated use of the product rule.

Exercise A.13 (B) If the elements of a determinant \( D(x) \) are differentiable functions of \( x \),
\[
D(x) = \begin{vmatrix}
    f(x) & g(x) \\
    \phi(x) & \psi(x)
\end{vmatrix}
\]
show that
\[
D'(x) = \begin{vmatrix}
    f'(x) & g'(x) \\
    \phi'(x) & \psi'(x)
\end{vmatrix} + \begin{vmatrix}
    f(x) & g(x) \\
    \phi'(x) & \psi'(x)
\end{vmatrix}.
\]
Extend this result to third-order determinants.

A.2.5 Mean Value Theorems

If a function \( f(x) \) is sufficiently smooth for all points inside the interval \( a < x < b \), its graph is a smooth curve\(^5\) starting at the point \( A = (a, f(a)) \) and ending at \( B = (b, f(b)) \), as shown in figure A.3.

From this figure it seems plausible that the tangent to the curve must be parallel to the chord \( AB \) at least once. That is
\[
f'(x) = \frac{f(b) - f(a)}{b - a} \quad \text{for some } x \text{ in the interval } a < x < b. \quad (A.8)
\]
Alternatively this may be written in the form
\[
f(b) = f(a) + hf'(a + \theta h), \quad h = b - a, \quad (A.9)
\]
Figure A.3: Diagram illustrating Cauchy’s form of the mean value theorem.

where \( \theta \) is a number in the interval \( 0 < \theta < 1 \), and is normally unknown. This relation is used frequently throughout the module. Note that equation (A.9) shows that between zeros of a continuous function there is at least one point at which the derivative is zero.

Equation (A.8) can be proved and is enshrined in the following theorem.

**Theorem A.1 The Mean Value Theorem (Cauchy’s form).** If \( f(x) \) and \( g(x) \) are real and differentiable for \( a \leq x \leq b \), then there is a point \( u \) inside the interval at which

\[
(f(b) - f(a))g'(u) = (g(b) - g(a))f'(u), \quad a < u < b.
\]

(A.10)

By putting \( g(x) = x \), equation (A.8) follows.

A similar idea may be applied to integrals. In figure A.4 is shown a typical continuous function, \( f(x) \), which attains its smallest and largest values, \( S \) and \( L \) respectively, on the interval \( a \leq x \leq b \).

Figure A.4: Diagram showing the upper and lower bounds of \( f(x) \) used to bound the integral.

It is clear that the area under the curve is greater than \( (b - a)S \) and less than \( (b - a)L \), that is

\[
(b - a)S \leq \int_{a}^{b} dx \ f(x) \leq (b - a)L.
\]

Because \( f(x) \) is continuous it follows that

\[
\int_{a}^{b} dx \ f(x) = (b - a)f(\xi) \quad \text{for some} \quad \xi \in [a, b].
\]

(A.11)

This observation is made rigorous in the following theorem.

**Theorem A.2 The Mean Value theorem (integral form).** If, on the closed interval \( a \leq x \leq b \), \( f(x) \) is continuous and \( \phi(x) \geq 0 \) then there is an \( \xi \) satisfying \( a \leq \xi \leq b \) such that

\[
\int_{a}^{b} dx \ f(x)\phi(x) = f(\xi) \int_{a}^{b} dx \ \phi(x).
\]

(A.12)
A.2. FUNCTIONS OF A REAL VARIABLE

If \( \phi(x) = 1 \) relation (A.11) is regained.

Exercise A.14 (B) The chain rule
In this exercise the Mean Value Theorem is used to derive the chain rule, equation (A.7), for the derivative of \( F(x) = f(g(x)) \).

Use the mean value theorem to show that

\[
F(x + h) - F(x) = f(g(x) + hg'(x + h\theta)) - f(g(x))
\]

and that

\[
f(g(x) + hg'(x + h\theta)) = f(g(x)) + hg'(x + h\theta)f'(g + h\phi g')
\]

where \( 0 < \theta, \phi < 1 \). Hence show that

\[
\frac{F(x + h) - F(x)}{h} = f'(g + h\phi g')g'(x + h\theta),
\]

and by taking the limit \( h \to 0 \) derive equation (A.7).

Exercise A.15 (B) Use the integral form of the mean value theorem, equation (A.11), to evaluate the limits,

(a) \( \lim_{x \to 0} \frac{1}{x} \int_0^x dt \sqrt{4 + 3t^2} \),
(b) \( \lim_{x \to 1} \frac{1}{(x - 1)^3} \int_1^x dt \ln (3t - 3t^2 + t^3) \).

A.2.6 Partial Derivatives

Here we consider functions of two or more variables, in order to introduce the idea of a partial derivative. If \( f(x, y) \) is a function of the two, independent variables \( x \) and \( y \), meaning that changes in one do not affect the other, then we may form the partial derivative of \( f(x, y) \) with respect to either \( x \) or \( y \) using a minor modification of the definition A.5 (page 354).

Definition A.6 The partial derivative of a function \( f(x, y) \) of two variables with respect to the first variable \( x \) is

\[
\frac{\partial f}{\partial x} = f_x(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}.
\]

In the computation of \( f_x \) the variable \( y \) is unchanged.

Similarly, the partial derivative with respect to the second variable \( y \) is

\[
\frac{\partial f}{\partial y} = f_y(x, y) = \lim_{k \to 0} \frac{f(x, y + k) - f(x, y)}{k}.
\]

In the computation of \( f_y \) the variable \( x \) is unchanged.

We use the conventional notation, \( \partial f/\partial x \), to denote the partial derivative with respect to \( x \), which is formed by fixing \( y \) and using the rules of ordinary calculus for the derivative with respect to \( x \). The suffix notation, \( f_x(x, y) \), is used to denote the same function: here the suffix \( x \) shows the variable being differentiated, and it has the advantage that when necessary it can be used in the form \( f_x(a, b) \) to indicate that the partial derivative \( f_x \) is being evaluated at the point \((a, b)\).
In practice the evaluation of partial derivatives is exactly the same as ordinary derivatives and the same rules apply. Thus if \( f(x, y) = xe^y \ln(2x + 3y) \) then the partial derivatives with respect to \( x \) and \( y \) are, respectively,

\[
\frac{\partial f}{\partial x} = e^y \ln(2x + 3y) + \frac{2xe^y}{2x + 3y} \quad \text{and} \quad \frac{\partial f}{\partial y} = xe^y \ln(2x + 3y) + \frac{3xe^y}{2x + 3y}.
\]

**Exercise A.16 (B)**

(a) If \( u = x^2 \sin(ln y) \) compute \( u_x \) and \( u_y \).

(b) If \( r^2 = x^2 + y^2 \) show that \( \frac{\partial r}{\partial x} = \frac{x}{r} \) and \( \frac{\partial r}{\partial y} = \frac{y}{r} \).

The partial derivatives are also functions of \( x \) and \( y \), so may be differentiated again. Thus we have

\[
\frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2} = f_{xx}(x, y) \quad \text{and} \quad \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial y^2} = f_{yy}(x, y). \quad (A.13)
\]

But now we also have the mixed derivatives

\[
\frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) \quad \text{and} \quad \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right). \quad (A.14)
\]

Except in special circumstances the order of differentiation is irrelevant so we obtain the mixed derivative rule

\[
\frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}. \quad (A.15)
\]

Using the suffix notation the mixed derivative rule is \( f_{xy} = f_{yx} \). A sufficient condition for this to hold is that both \( f_{xy} \) and \( f_{yx} \) are continuous functions of \( (x, y) \), see equation (A.4) (page 352).

Similarly, differentiating \( p \) times with respect to \( x \) and \( q \) times with respect to \( y \), in any order, gives the same \( n \)th order derivative,

\[
\frac{\partial^n f}{\partial x^p \partial y^q} \quad \text{where} \quad n = p + q,
\]

provided all the \( n \)th derivatives are continuous.

**Exercise A.17 (B)** If \( \Phi(x, y) = \exp(-x^2/y) \) show that \( \Phi \) satisfies the equations

\[
\frac{\partial \Phi}{\partial x} = -\frac{2x \Phi}{y} \quad \text{and} \quad \frac{\partial^2 \Phi}{\partial x^2} = 4\frac{\partial \Phi}{\partial y} - \frac{2\Phi}{y}. \quad \Box
\]

**Exercise A.18 (B)** Show that \( u = x^2 \sin(ln y) \) satisfies the equation \( 2y^2 \frac{\partial^2 u}{\partial y^2} + 2y \frac{\partial u}{\partial y} + x \frac{\partial u}{\partial x} = 0 \). \quad \Box

The generalisation of these ideas to functions of the \( n \) variables \( x = (x_1, x_2, \ldots, x_n) \) is straightforward: the partial derivative of \( f(x) \) with respect to \( x_k \) is defined to be

\[
\frac{\partial f}{\partial x_k} = \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_{k-1}, x_k + h, x_{k+1}, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{h}. \quad (A.16)
\]
A.2. FUNCTIONS OF A REAL VARIABLE

All other properties of the derivatives are the same as in the case of two variables, in particular for the mth derivative the order of differentiation is immaterial provided all mth derivatives are continuous.

For a function of a single variable, $f(x)$, the existence of the derivative, $f'(x)$, implies that $f(x)$ is continuous. For functions of two or more variables the existence of the partial derivatives does not guarantee continuity.

The total derivative

If $f(x_1, x_2, \ldots, x_n)$ is a function of $n$ variables and if each of these variables is a function of the single variable $t$, we may form a new function of $t$ with the formula

$$F(t) = f(x_1(t), x_2(t), \ldots, x_n(t)). \quad (A.17)$$

Geometrically, $F(t)$ represents the value of $f(x)$ on a curve $C$ defined parametrically by the functions $(x_1(t), x_2(t), \ldots, x_n(t))$. The derivative of $F(t)$ is given by the relation

$$\frac{dF}{dt} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \frac{dx_k}{dt} \quad (A.18)$$

so $F'(t)$ is the rate of change of $f(x)$ along $C$. Normally, we write $f(t)$ rather than use a different symbol $F(t)$, and the left-hand side of the above equation is written $\frac{df}{dt}$.

This derivative is named the total derivative of $f$. The proof of this when $n = 2$ and $x'$ and $y'$ do not vanish near $(x, y)$ is sketched below; the generalisation to larger $n$ is straightforward. If $F(t) = f(x(t), y(t))$ then

$$F(t + \epsilon) = f(x(t + \epsilon), y(t + \epsilon)) = f(x(t) + \epsilon x'(t + \theta \epsilon), y(t) + \epsilon y'(t + \phi \epsilon)), \quad 0 < \theta, \phi < 1,$$

where we have used the mean value theorem, equation (A.9). Write the right-hand side in the form

$$f(x + \epsilon x', y + \epsilon y') = [f(x + \epsilon x', y + \epsilon y') - f(x, y + \epsilon y')] + [f(x, y + \epsilon y') - f(x, y)] + f(t)$$

so that

$$\frac{F(t + \epsilon) - F(t)}{\epsilon} = \frac{f(x + \epsilon x', y + \epsilon y') - f(x, y + \epsilon y')}{\epsilon x'} x' + \frac{f(x, y + \epsilon y') - f(x, y)}{\epsilon y'} y'.$$

Thus, on taking the limit as $\epsilon \to 0$ we have

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

This result remains true if either or both $x' = 0$ or $y' = 0$, but then more care is needed with the proof.

Equation (A.18) is used in chapter 4 to derive one of the most important results in the module: if the dependence of $x$ upon $t$ is linear and $F(t)$ has the form

$$F(t) = f(x + th) = f(x_1 + th_1, x_2 + th_2, \ldots, x_n + th_n)$$
where the vector $h$ is constant and the variable $x_k$ has been replaced by $x_k + t h_k$, for all $k$. Since $\frac{d}{dt}(x_k + th_k) = h_k$, equation (A.18) becomes

$$\frac{dF}{dt} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} h_k.$$  \hspace{1cm} (A.19)

This result will also be used in section A.2.9 to derive the Taylor series for several variables.

A variant of equation (A.17), which frequently occurs in the Calculus of Variations, is the case where $f(x)$ depends explicitly upon the variable $t$, so this equation becomes

$$F(t) = f(t, x_1(t), x_2(t), \ldots, x_n(t))$$

and then equation (A.18) acquires an additional term,

$$\frac{dF}{dt} = \frac{\partial f}{\partial t} + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \frac{dx_k}{dt}.$$  \hspace{1cm} (A.20)

For an example we apply this formula to the function

$$f(t, x, y) = x \sin(yt) \text{ with } x = e^t \text{ and } y = e^{-2t},$$

so

$$F(t) = f(t, e^t, e^{-2t}) = e^t \sin(te^{-2t}).$$

Equation (A.20) gives

$$\frac{dF}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

$$= xy \cos(yt) + e^t \sin(yt) - 2xt \cos(yt)e^{-2t},$$

which can be expressed in terms of $t$ only,

$$\frac{dF}{dt} = (1 - 2t)e^{-t} \cos(te^{-2t}) + e^t \sin(te^{-2t}).$$

The same expression can also be obtained by direct differentiation of $F(t) = e^t \sin(te^{-2t})$. The right-hand sides of equations (A.18) and (A.20) depend upon both $x$ and $t$, but because $x$ depends upon $t$ often these expressions are written in terms of $t$ only. In the Calculus of Variations this is usually not helpful because the dependence of both $x$ and $t$, separately, is important: for instance we often require expressions like

$$\frac{d}{dt} \left( \frac{\partial F}{\partial x_1} \right) \text{ and } \frac{\partial}{\partial x_1} \left( \frac{dF}{dt} \right).$$

The second of these expressions requires some clarification because $dF/dt$ contains the derivatives $x'_k$. Thus

$$\frac{\partial}{\partial x_1} \left( \frac{dF}{dt} \right) = \frac{\partial}{\partial x_1} \left( \frac{\partial f}{\partial t} + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \frac{dx_k}{dt} \right).$$
Since \( x_k'(t) \) is independent of \( x_1 \) for all \( k \), this becomes

\[
\frac{\partial}{\partial x_1} \left( \frac{dF}{dt} \right) = \frac{\partial^2 f}{\partial x_1 \partial t} + \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_1 \partial x_k} \frac{dx_k}{dt} = \frac{d}{dt} \left( \frac{\partial F}{\partial x_1} \right),
\]

the last line being a consequence of the mixed derivative rule.

**Exercise A.19 (B)** If \( f(t, x, y) = xy - ty^2 \) and \( x = t^2, \ y = t^3 \) show that

\[
\frac{df}{dt} = -y^2 + y \frac{dx}{dt} + \frac{dy}{dt} (x - 2ty) = t^4(5 - 7t^2),
\]

and that

\[
\frac{\partial}{\partial y} \left( \frac{df}{dt} \right) = \frac{dx}{dt} - 2y - 2t \frac{dy}{dt} = 2t (1 - 4t^2),
\]

\[
\frac{d}{dt} \left( \frac{\partial f}{\partial y} \right) = \frac{d}{dt} (x - 2ty) = \frac{dx}{dt} - 2y - 2t \frac{dy}{dt} = 2t (1 - 4t^2).
\]

**Exercise A.20 (B)** If \( F = \sqrt{1+x_1x_2} \), and \( x_1 \) and \( x_2 \) are functions of \( t \), show by direct calculation of each expression that

\[
\frac{\partial}{\partial x_1} \left( \frac{dF}{dt} \right) = \frac{d}{dt} \left( \frac{\partial F}{\partial x_1} \right) = \frac{x_2'}{2\sqrt{1+x_1x_2}} - \frac{x_2(x_1'x_2 + x_1x_2')}{4(1+x_1x_2)^{3/2}}.
\]

**Exercise A.21 (B)** Euler’s formula for homogeneous functions

(a) A function \( f(x, y) \) is said to be homogeneous with degree \( p \) in \( x \) and \( y \) if it has the property \( f(\lambda x, \lambda y) = \lambda^p f(x, y) \), for some fixed real number \( p \) and for any constant \( \lambda \). For such a function prove Euler’s formula:

\[
pf(x, y) = x f_x(x, y) + y f_y(x, y).
\]

Hint use the total derivative formula (A.18) and differentiate with respect to \( \lambda \).

(b) Find the equivalent result for homogeneous functions of \( n \) variables that satisfy \( f(\lambda x) = \lambda^p f(x) \).

(c) Show that if \( f(x_1, x_2, \ldots, x_n) \) is a homogeneous function of degree \( p \), then each of the partial derivatives, \( \partial f/\partial x_k \), \( k = 1, 2, \ldots, n \), is a homogeneous function of degree \( p - 1 \).

**A.2.7 Implicit functions**

An equation of the form \( f(x, y) = 0 \), where \( f \) is a suitably well behaved function of both \( x \) and \( y \), can define a curve in the Cartesian plane, as illustrated in figure A.5.

For some values of \( x \) the equation \( f(x, y) = 0 \) can be solved to yield one or more real values of \( y \), which will give one or more functions of \( x \). For instance the equation
\[ x^2 + y^2 - 1 = 0 \] defines a circle in the plane and for each \( x \) in \( |x| < 1 \) there are two values of \( y \), giving the two functions \( y(x) = \pm \sqrt{1 - x^2} \). A more complicated example is the equation \( x - y + \sin(xy) = 0 \), which cannot be rearranged to express one variable in terms of the other.

Consider the smooth curve sketched in figure A.5. On a segment in which the curve is not parallel to the \( y \)-axis the equation \( f(x, y) = 0 \) defines a function \( y(x) \). Such a function is said to be defined implicitly. The same equation will also define \( x(y) \), that is \( x \) as a function of \( y \), provided the segment does not contain a point where the curve is parallel to the \( x \)-axis. This result, inferred from the picture, is a simple example of the implicit function theorem stated below.

Implicitly defined functions are important because they occur frequently as solutions of differential equations, see exercise A.25, but there are few, if any, general rules that help understand them. It is, however, possible to obtain relatively simple expressions for the first derivatives, \( y'(x) \) and \( x'(y) \).

We assume that \( y(x) \) exists and is differentiable, as seems reasonable from figure A.5, so \( F(x) = f(x, y(x)) \) is a function of \( x \) only and we may use the chain rule (A.20) to differentiate with respect to \( x \). This gives

\[
\frac{dF}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx}.
\]

On the curve defined by \( f(x, y) = 0 \), \( F'(x) = 0 \) and hence

\[
\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx} = 0 \quad \text{or} \quad \frac{dy}{dx} = \frac{-f_x}{f_y}.
\]  \hspace{1cm} (A.21)

Similarly, if \( x(y) \) exists and is differentiable a similar analysis using \( y \) as the independent variable gives

\[
\frac{\partial f}{\partial x} \frac{dx}{dy} + \frac{\partial f}{\partial y} = 0 \quad \text{or} \quad \frac{dx}{dy} = \frac{-f_y}{f_x}.
\]  \hspace{1cm} (A.22)

This result is encapsulated in the Implicit Function Theorem which gives sufficient conditions for an equation of the form \( f(x, y) = 0 \) to have a ‘solution’ \( y(x) \) satisfying \( f(x, y(x)) = 0 \). A restricted version of it is given here.

**Theorem A.3 Implicit Function Theorem:** Suppose that \( f : U \to R \) is a function with continuous partial derivatives defined in an open set \( U \subseteq \mathbb{R}^2 \). If there is a point \((a, b) \in U\) for which \( f(a, b) = 0 \) and \( f_y(a, b) \neq 0 \), then there are open intervals \( I = (x_1, x_2) \) and \( J = (y_1, y_2) \) such that \((a, b)\) lies in the rectangle \( I \times J \) and for every \( x \in I \) \( f(x, y) = 0 \) determines exactly one value \( y(x) \in J \) for which \( f(x, y(x)) = 0 \).
A.2. FUNCTIONS OF A REAL VARIABLE

The function $y : I \to J$ is continuous, differentiable, with the derivative given by equation (A.21).

Exercise A.22 (B) In the case $f(x, y) = y - g(x)$ show that equations (A.21) and (A.22) leads to the relation

$$\frac{dx}{dy} = \left(\frac{dy}{dx}\right)^{-1}.$$ 

Exercise A.23 (B) If $\ln(x^2 + y^2) = 2 \tan^{-1}(y/x)$ find $y'(x)$. 

Exercise A.24 (B,E) If $x - y + \sin(xy) = 0$ determine the values of $y'(0)$ and $y''(0)$. 

Exercise A.25 (B) Show that the differential equation

$$\frac{dy}{dx} = \frac{y - a^2 x}{y + x}, \quad y(1) = A > 0,$$

has a solution defined by the equation

$$\frac{1}{2} \ln \left( a^2 x^2 + y^2 \right) + \frac{1}{a} \tan^{-1} \left( \frac{y}{ax} \right) = B \quad \text{where} \quad B = \frac{1}{2} \ln \left( a^2 + A^2 \right) + \frac{1}{a} \tan^{-1} \left( \frac{A}{a} \right).$$

Hint the equation may be put in separable form by defining a new dependent variable $v = y/x$. 

The implicit function theorem can be generalised to deal with the set of functions

$$f_k(x, t) = 0, \quad k = 1, 2, \ldots, n,$$

where $x = (x_1, x_2, \ldots, x_n)$ and $t = (t_1, t_2, \ldots, t_n)$. These $n$ equations have a unique solution for each $x_k$ in terms of $t$, $x_k = g_k(t)$, $k = 1, 2, \ldots, n$, in the neighbourhood of $(x_0, t_0)$ provided that at this point the derivatives $\partial f_j/\partial x_k$, exist and that the determinant

$$J = \left| \begin{array}{ccc}
\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_n} \\
\frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_n} \\
\vdots & \vdots & \vdots \\
\frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \frac{\partial f_n}{\partial x_n}
\end{array} \right| (A.24)$$

is not zero. Furthermore all the functions $g_k(t)$ have continuous first derivatives. The determinant $J$ is named the Jacobian determinant or, more usually, the Jacobian. It is often helpful to use either of the following notations for the Jacobian,

$$J = \frac{\partial f}{\partial x} \quad \text{or} \quad J = \frac{\partial (f_1, f_2, \ldots, f_n)}{\partial (x_1, x_2, \ldots, x_n)} (A.25)$$

Exercise A.26 (B) Show that the equations $x = r \cos \theta$, $y = r \sin \theta$ can be inverted to give functions $r(x, y)$ and $\theta(x, y)$ in every open set of the plane that does not include the origin.
A.2.8 Taylor series for one variable

The Taylor series is a method of representing a given sufficiently well behaved function in terms of an infinite power series, defined in the following theorem.

**Theorem A.4 Taylor’s Theorem:** If \( f(x) \) is a function defined on \( x_1 \leq x \leq x_2 \) such that \( f^{(n)}(x) \) is continuous for \( x_1 \leq x \leq x_2 \) and \( f^{(n+1)}(x) \) exists for \( x_1 < x < x_2 \), then if \( a \in [x_1,x_2] \) for every \( x \in [x_1,x_2] \)

\[
f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a) + R_{n+1}. \tag{A.26}
\]

The remainder term, \( R_{n+1} \), can be expressed in the form

\[
R_{n+1} = \frac{(x-a)^{n+1}}{(n+1)!}f^{(n+1)}(a + \theta h) \quad \text{for some} \quad 0 < \theta < 1 \quad \text{and} \quad h = x - a. \tag{A.27}
\]

If all derivatives of \( f(x) \) are continuous for \( x_1 \leq x \leq x_2 \), and if the remainder term \( R_n \to 0 \) as \( n \to \infty \) in a suitable manner, we may take the limit to obtain the infinite series

\[
f(x) = \sum_{k=0}^{\infty} \frac{(x-a)^k}{k!}f^{(k)}(a). \tag{A.28}
\]

The infinite series (A.28) is known as Taylor’s series, and the point \( x = a \) the point of expansion. A similar series exists when \( x \) takes complex values.

Care is needed when taking the limit of (A.26) as \( n \to \infty \), because there are cases when the infinite series on the right-hand side of equation (A.28) does not equal \( f(x) \). If, however, the Taylor series converges to \( f(x) \) at \( x = \xi \) then for any \( x \) closer to \( a \) than \( \xi \), that is \( |x-a| < |\xi-a| \), the series converges to \( f(x) \). This caveat is necessary because of the strange example \( g(x) = \exp(-1/x^2) \) for which all derivatives are continuous and are zero at \( x = 0 \); for this function the Taylor series about \( x = 0 \) can be shown to exist, but for all \( x \) it converges to zero rather than \( g(x) \). This means that for any well behaved function, \( f(x) \) say, with a Taylor series that converges to \( f(x) \) at a different function, \( f(x) + g(x) \) can be formed whose Taylor series converges, but to \( f(x) \) not \( f(x) + g(x) \). This strange behaviour is not uncommon in functions arising from physical problems; however, it is ignored in this module and we shall assume that the Taylor series derived from a function converges to it in some interval.

The series (A.28) was first published by Brook Taylor (1685–1731) in 1715: the result obtained by putting \( a = 0 \) was discovered by Stirling (1692–1770) in 1717 but first published by Maclaurin (1698–1746) in 1742. With \( a = 0 \) this series is therefore often known as Maclaurin’s series.

In practice, of course, it is usually impossible to sum the infinite series (A.28), so it is necessary to truncate it at some convenient point and this requires knowledge of how, or indeed whether, the series converges to the required value. Truncation gives rise to the Taylor polynomials, with the order-\( n \) polynomial given by

\[
f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!}f^{(k)}(a). \tag{A.29}
\]
A.2. FUNCTIONS OF A REAL VARIABLE

The series (A.28) is an infinite series of the functions $(x - a)^n f^{(n)}(a)/n!$ and summing these requires care. A proper understanding of this process requires careful definitions of convergence which may be found in any text book on analysis. For our purposes, however, it is sufficient to note that in most cases there is a real number, $r_c$, named the radius of convergence, such that if $|x - a| < r_c$ the infinite series is well mannered and behaves rather like a finite sum: the value of $r_c$ can be infinite, in which case the series converges for all $x$.

If the Taylor series of $f(x)$ and $g(x)$ have radii of convergence $r_f$ and $r_g$ respectively, then the Taylor series of $\alpha f(x) + \beta g(x)$, for constants $\alpha$ and $\beta$, and of $f(x)^n g(x)^k$, for positive constants $\alpha$ and $\beta$, exist and have the radius of convergence $\min(r_f, r_g)$. The Taylor series of the compositions $f(g(x))$ and $g(f(x))$ may also exist, but their radii of convergence depend upon the behaviour of $g$ and $f$ respectively. Also Taylor series may be integrated and differentiated to give the Taylor series of the integral and derivative of the original function, and with the same radius of convergence.

Formally, the $n$th Taylor polynomial of a function is formed from its first $n$ derivatives at the point of expansion. In practice, however, the calculation of high-order derivatives is very awkward and it is often easier to proceed by other means, which rely upon ingenuity. A simple example is the Taylor series of $\ln(1 + \tanh x)$, to fourth order; this is most easily obtained using the known Taylor expansions of $\ln(1 + z)$ and $\tanh x$,

$$\ln (1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + O(z^5)$$

and $\tanh x = x - \frac{x^3}{3} + \frac{2x^5}{15} + O(x^7)$,

and then put $z = \tanh x$ retaining only the appropriate order of the series expansion. Thus

$$\ln(1 + \tanh x) = \left[ x - \frac{x^3}{3} + O(x^5) \right] - \left[ \frac{x^2}{2} \left( 1 - \frac{x^2}{3} + \ldots \right)^2 \right] + \frac{x^3}{3} - \frac{x^4}{4} + O(x^5)$$

$$= x - \frac{x^2}{2} + \frac{x^4}{12} + O(x^5).$$

This method is far easier than computing the four required derivatives of the original function.

For $|x - a| > r_c$ the infinite sum (A.28) does not exist. It follows that knowledge of $r_c$ is important. It can be shown that, in most cases of practical interest, its value is given by either of the limits

$$r_c = \lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right| \quad \text{or} \quad r_c = \lim_{n \to \infty} |a_n|^{-1/n} \quad \text{where} \quad a_k = \frac{f^{(k)}(a)}{k!}. \quad (A.30)$$

Usually the first expression is most useful. Typically, we have, for large $n$

$$\left| \frac{n!}{f^{(n)}(a)} \right|^{1/n} = r_c (1 + O(1/n))$$

so that

$$\frac{n!}{f^{(n)}(a)} = \frac{A r_c^n}{(1 + O(1/n))}$$

for some constant $A$. Then the $n$th term of the series behaves as $((x - a)/r_c)^n$, and decreases rapidly with increasing $n$ provided $|x - a| < r_c$ and $n$ is sufficiently large.

Superficially, the Taylor series appears to be a useful representation and a good approximation. In general this is not true unless $|x - a|$ is small; for practical applications far more efficient approximations exist — that is they achieve the same accuracy for
far less work. The basic problem is that the Taylor expansion uses knowledge of the function at one point only, and the larger \(|x - a|\) the more terms are required for a given accuracy. More sensible approximations, on a given interval, take into account information from the whole interval: we describe some approximations of this type in chapter 13.

The first practical problem is that the remainder term, equation (A.27), depends upon \(\theta\), the value of which is unknown. Hence \(R_n\) cannot be computed; also, it is normally difficult to estimate.

In order to understand how these series converge we need to consider the magnitude of the \(n\)th term in the Taylor series: this type of analysis is important for any numerical evaluation of power series. The \(n\)th term is a product of \((x - a)^n/n!\) and \(f^{(n)}(a)\). Using Stirling’s approximation,

\[
n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n (1 + O(1/n)) \tag{A.31}\]

we can approximate the first part of this product by

\[
\left|\frac{(x - a)^n}{n!}\right| \simeq \frac{1}{\sqrt{2\pi n}} \left(\frac{e|x - a|}{n}\right)^n = g_n. \tag{A.32}\]

The expression \(g_n\) decreases very rapidly with increasing \(n\), provided \(n\) is large enough. Hence the term \(|x - a|^n/n!\) may be made as small as we please. But for practical applications this is not sufficient; in figure A.6 we plot a graph of the values of \(\log(g_n)\), that is the logarithm to the base 10, for \(x - a = 10\).

![Figure A.6: Graph showing the value of \(\log(g_n)\), equation (A.32), for \(x - a = 10\). For clarity we have joined the points with a continuous line.](image)

In this example the maximum of \(g_n\) is at \(n = 10\) and has a value of about 2500, before it starts to decrease. It is fairly simple to show that \(g_n\) has a maximum at \(n \simeq |x - a|\) and here its value is \(\max(g_n) \simeq \exp(|x - a|)/\sqrt{2\pi|x - a|}\).

The value of \(f^{(n)}(a)\) is also difficult to estimate, but it usually increases rapidly with \(n\). Bizarrely, in many cases of interest, this behaviour depends upon the behaviour of \(f(z)\), where \(z\) is a complex variable. An understanding of this requires a study of Complex Variable Theory, which is beyond the scope of this chapter. Instead we illustrate the behaviour of Taylor polynomials with a simple example.

First consider the Taylor series of \(\sin x\), about \(x = 0\),

\[
\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!} + \ldots, \tag{A.33}\]
which is derived in exercise A.27.

Note that only odd powers occur, because \( \sin x \) is an odd function, and also that the radius of convergence is infinite. In figure A.7 we show graphs of this series, truncated at \( x^{2n-1} \) with \( n = 1, 4, 8 \) and 15 for \( 0 < x < 4\pi \).

![Figure A.7: Graph comparing the Taylor polynomials, of order \( n \), for the sine function with the exact function, the dashed line.](image)

These graphs show that for large \( x \) it is necessary to include many terms in the series to obtain an accurate representation of \( \sin x \). The reason is simply that for fixed, large \( x \), \( x^{2n-1}/(2n-1)! \) is very large at \( n = x \), as shown in figure A.6. Because the terms of this series alternate in sign the large terms in the early part of the series partially cancel and cause problems when approximating a function \( O(1) \): it is worth noting that as a consequence, with a computer having finite accuracy there is a value of \( x \) beyond which the Taylor series for \( \sin x \) gives incorrect values, despite the fact that formally it converges for all \( x \).

**Exercise A.27 (B) Exponential and Trigonometric functions**

If \( f(x) = \exp(ix) \) show that \( f^{(n)}(x) = i^n \exp(ix) \) and hence that its Taylor series is

\[
e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!}.
\]

Show that the radius of convergence of this series in infinite. Deduce that

\[
\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{(-1)^n x^{2n}}{(2n)!} + \cdots,
\]

\[
\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + \cdots.
\]

**Exercise A.28 (B) Binomial expansion**

Show that the Taylor series of \( (1 + x)^a \) is

\[
(1 + x)^a = 1 + ax + \frac{1}{2}a(a-1)x^2 + \cdots + \frac{a(a-1)(a-2)\cdots(a-k+1)}{k!}x^k + \cdots.
\]

When \( a = n \) is an integer this series terminates at \( k = n \) and becomes the binomial expansion

\[
(1 + x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \quad \text{where} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}.
\]
are the binomial coefficients.

Exercise A.29 (B) If \( f(x) = \tan x \) find the first three derivatives to show that 
\[ \tan x = x + \frac{1}{3} x^3 + O(x^5). \]

Exercise A.30 (B) The natural logarithm

(a) Show that 
\[ \frac{1}{1 + t} = 1 - t + t^2 + \cdots + (-1)^n t^n + \cdots \]
and use the definition of the natural logarithm, \( \ln(1 + x) = \int_0^x \frac{dt}{1 + t} \) to show that 
\[ \ln (1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{(-1)^{n-1} x^n}{n} + \cdots. \]

(b) For which values of \( x \) is this expression valid?

(c) Use this result to show that 
\[ \ln \left( \frac{1 + x}{1 - x} \right) = 2 \left( x + \frac{x^3}{3} + \cdots + \frac{x^{2n-1}}{2n-1} + \cdots \right). \]

Exercise A.31 (B) The inverse tangent function

Use the definition 
\[ \tan^{-1} x = \int_0^x \frac{dt}{1 + t^2} \]
to show that for \( |x| < 1, \)
\[ \tan^{-1} x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1}. \]

Exercise A.32 (B) Show that 
\[ \ln(1 + \sinh x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{5x^4}{12} + O(x^5). \]

Exercise A.33 (B) Obtain the first five terms of the Taylor series of the function that satisfies the equation 
\[ (1 + x) \frac{dy}{dx} = 1 + xy + y^2, \quad y(0) = 0. \]

Hint use Leibniz’s rule given in exercise A.11 (page 357) to differentiate the equation \( n \) times.

A.2.9 Taylor series for several variables

The Taylor series of a function \( f : \mathbb{R}^n \to \mathbb{R} \) is trivially derived from the Taylor expansion of a function of one variable using the chain rule, equation (A.19) (page 362). The only difficulty is that the algebra very quickly becomes unwieldy with increasing order.

We require the expansion of \( f(x) \) about \( x = a \), so we need to represent \( f(a + h) \) as some sort of power series in \( h \). To this end, define a function of the single variable \( t \) by the relation 
\[ F(t) = f(a + th) \quad \text{so} \quad F(0) = f(a), \]
and \( F(t) \) gives values of \( f(x) \) on the straight line joining \( a \) to \( a + h \). The Taylor series of \( F(t) \) about \( t = 0 \) is, on using equation (A.26) (page 366),

\[
F(t) = F(0) + tF'(0) + \frac{t^2}{2} F''(0) + \cdots + \frac{t^n}{n!} F^{(n)}(0) + R_{n+1},
\]

which we assume to exist for \(|t| \leq 1\). Now we need only express the derivatives \( F^{(n)}(0) \) in terms of the partial derivatives of \( f(x) \). Equation (A.19) (page 362) gives

\[
F'(0) = \sum_{k=1}^{m} f_{x_k} (a) h_k.
\]

Hence to first order the Taylor series is

\[
f(a + h) = f(a) + \sum_{k=1}^{m} h_k f_{x_k} (a) + R_2 = f(a) + h \cdot \frac{\partial f}{\partial a} + R_2,
\]

where \( R_2 \) is the remainder term which is second order in \( h \) and is given below. Here we have introduced the notation \( \partial f / \partial x \) for the vector function,

\[
\frac{\partial f}{\partial x} = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m} \right)
\]

with the scalar product \( h \cdot \frac{\partial f}{\partial x} = \sum_{k=1}^{m} h_k \frac{\partial f}{\partial x_k} \).

For the second derivative we use equation (A.19) (page 362) again,

\[
F''(t) = \sum_{k=1}^{m} h_k \frac{d}{dt} f_{x_k} (a + th) = \sum_{k=1}^{m} h_k \left( \sum_{i=1}^{m} h_i f_{x_i, x_k} (a + th) \right).
\]

At \( t = 0 \) this can be written in the form,

\[
F''(0) = \sum_{k=1}^{m} h_k \sum_{i=1}^{m} h_i f_{x_i, x_k} (a)
\]

\[
= \sum_{k=1}^{m} h_k^2 f_{x_k, x_k} (a) + 2 \sum_{k=1}^{m-1} \sum_{i=k+1}^{m} h_k h_i f_{x_i, x_k} (a),
\]

where the second relation comprises fewer terms because the mixed derivative rule has been used. This gives the second order Taylor series,

\[
f(a + h) = f(a) + \sum_{k=1}^{m} h_k f_{x_k} (a) + \frac{1}{2!} \left( \sum_{k=1}^{m} h_k \sum_{i=1}^{m} h_i f_{x_i, x_k} (a) \right) + R_3,
\]

where the remainder term is given below.

The higher-order terms are derived in exactly the same manner, but the algebra quickly becomes cumbersome. It helps, however, to use the linear differential operator \( h \cdot \partial / \partial a \) to write the derivatives of \( F(t) \) at \( t = 0 \) in the more convenient form,

\[
F'(0) = \left(h \cdot \frac{\partial}{\partial a}\right) f(a), \quad F''(0) = \left(h \cdot \frac{\partial}{\partial a}\right)^2 f(a), \quad F^{(n)}(0) = \left(h \cdot \frac{\partial}{\partial a}\right)^n f(a).
\]
Then we can write Taylor series in the form

\[ f(a + h) = f(a) + \sum_{s=1}^{n} \frac{1}{s!} \left( h \cdot \frac{\partial}{\partial a} \right)^s f(a) + R_{n+1} \quad (A.39) \]

where the remainder term is

\[ R_{n+1} = \frac{1}{(n + 1)!} f^{(n+1)}(\theta) \text{ for some } 0 < \theta < 1. \]

Because the high order derivatives are so cumbersome and for the practical reasons discussed in section A.2.8, in particular figure A.7 (page 369), Taylor series for many variables are rarely used beyond the second order term. This term, however, is important for the classification of stationary points, considered in chapter 8.

For functions of two variables, \((x, y)\), the Taylor series is

\[ f(a + h, b + k) = f(a, b) + hf_x + kf_y + \frac{1}{2} \left( h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy} \right) \]
\[ + \frac{1}{6} \left( h^3 f_{xxx} + 3h^2 kf_{xyy} + 3hk^2 f_{xxy} + 3k^3 f_{yyy} \right) + \ldots \]
\[ + \sum_{r=0}^{s} \frac{h^{s-r} k^r}{(s - r)!r!} \frac{\partial^s f}{\partial x^{s-r} \partial y^r} + \cdots + R_{n+1}, \quad (A.40) \]

where all derivatives are evaluated at \((a, b)\). In this case the \(s\)th term is relatively easy to obtain by expanding the differential operator \((h\partial/\partial x + k\partial/\partial y)^s\) using the binomial expansion (which works because the mixed derivative rule means that the two operators \(\partial/\partial x\) and \(\partial/\partial y\) commute).

**Exercise A.34 (B,E)** Find the Taylor expansions about \(x = y = 0\), up to and including the second order terms, of the functions

(a) \(f(x, y) = \sin x \sin y\), \hspace{1cm} (b) \(f(x, y) = \sin(x + e^{-y} - 1)\).

**Exercise A.35 (B)** Show that the third-order Taylor series for a function, \(f(x, y, z)\), of three variables is

\[ f(a + h, b + k, c + l) = f(a, b, c) + hf_x + kf_y + lf_z \]
\[ + \frac{1}{2!} \left( h^2 f_{xx} + k^2 f_{yy} + l^2 f_{zz} + 2hk f_{xy} + 2kl f_{yz} + 2lh f_{zx} \right) \]
\[ + \frac{1}{3!} \left( h^3 f_{xxx} + k^3 f_{yyy} + l^3 f_{zzz} + 6hk^2 f_{xxy} + 6hl^2 f_{yxz} + 6kl^2 f_{yzx} \right) \]
\[ + 3h^2 k^2 f_{xyy} + 3h^2 f_{xxx} + 3k^2 f_{yyy} + 3l^2 f_{zzz} + \cdots + R_{n+1}. \]

**A.2.10 L’Hospital’s rule**

Ratios of functions occur frequently and if

\[ R(x) = \frac{f(x)}{g(x)} \quad (A.41) \]
the value of $R(x)$ is normally computed by dividing the value of $f(x)$ by the value of $g(x)$: this works provided $g(x)$ is not zero at the point in question, $x = a$ say. If $g(x)$ and $f(x)$ are simultaneously zero at $x = a$, the value of $R(a)$ may be redefined as a limit. For instance if
\[ R(x) = \frac{\sin x}{x} \]  
(A.42)
then the value of $R(0)$ is not defined, though $R(x)$ does tend to the limit $R(x) \to 1$ as $x \to 0$. Here we show how this limit may be computed using L'Hospital’s rule\(^6\) and its extensions, discovered by the French mathematician G F A Marquis de l’Hospital (1661–1704).

Suppose that at $x = a$, $f(a) = g(a) = 0$ and that each function has a Taylor series about $x = a$, with finite radii of convergence: thus near $x = a$ we have for small, non-zero $|\epsilon|$, 
\[ R(a + \epsilon) = \frac{f(a + \epsilon)}{g(a + \epsilon)} = \frac{\epsilon f'(a) + O(\epsilon^2)}{\epsilon g'(a) + O(\epsilon^2)} = \frac{f'(a)}{g'(a)} + O(\epsilon) \quad \text{provided} \quad g'(a) \neq 0. \]
Hence, on taking the limit $\epsilon \to 0$, we obtain the result given by the following theorem.

**Theorem A.5 L’Hospital’s rule.** Suppose that $f(x)$ and $g(x)$ are real and differentiable for $-\infty < a < x < b < \infty$. If
\[
\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0 \quad \text{or} \quad \lim_{x \to a} f(x) = \pm \infty \quad \text{and} \quad \lim_{x \to a} g(x) = \pm \infty
\]
then
\[
\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}, \quad (A.43)
\]
provided the right-hand limit exists.

More generally if $f^{(k)}(a) = g^{(k)}(a) = 0$, $k = 0, 1, \ldots, n - 1$ and $g^{(n)}(a) \neq 0$ then
\[
\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f^{(n)}(x)}{g^{(n)}(x)},
\]
provided the right-hand limit exists.

Consider the function defined by equation (A.42); at $x = 0$ L’Hospital’s rule gives
\[
R(0) = \lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\cos x}{1} = 1.
\]

**Exercise A.36 (B)** Find the values of the following limits:

(a) $\lim_{x \to a} \frac{\cosh x - \cosh a}{\sinh x - \sinh a}$, \hspace{1cm} (b) $\lim_{x \to 0} \frac{\sin x - x}{x \cos x - x}$, \hspace{1cm} (c) $\lim_{x \to 0} \frac{3^x - 3^{-x}}{2x - 2^{-x}}$ \hspace{1cm} $\square$

**Exercise A.37 (B)**

(a) If $f(a) = g(a) = 0$ and $\lim_{x \to a} \frac{f'(x)}{g'(x)} = \infty$ show that $\lim_{x \to a} \frac{f(x)}{g(x)} = \infty$. 

(b) If both $f(x)$ and $g(x)$ are positive in a neighbourhood of $x = a$, tend to infinity as $x \to a$ and $\lim_{x \to a} \frac{f(x)}{g(x)} = A$ show that $\lim_{x \to a} \frac{f(x)}{g(x)} = A$. 

\(^6\)Here we use the spelling of the French national bibliography, as used by L’Hospital. Some modern text use the spelling L’Hôpital, instead of the silent s.
A.2.11 Integration

The study of integration arose from the need to compute areas and volumes. The theory of integration was developed independently from the theory of differentiation and the Fundamental Theorem of Calculus, described in note P:I on page 375 relates these processes. It should be noted, however, that Newton knew of the relation between gradients and areas and exploited it in his development of the subject.

In this section we provide a very brief outline of the simple theory of integration and discuss some of the methods used to evaluate integrals. This section is included for reference purposes; however, although the theory of integration is not central to the main topic of this module, you should be familiar with its contents. The important idea, needed in chapter 4, is that of differentiating with respect to a parameter, or ‘differentiating under the integral sign’ described in equation (A.50) (page 378).

In this discussion of integration we use an intuitive notion of area and refer the reader to suitable texts, Apostol (1963), Rudin (1976) or Whittaker and Watson (1965) for instance, for a rigorous treatment.

If \( f(x) \) is a real, continuous function of the interval \( a \leq x \leq b \), it is intuitively clear that the area between the graph and the \( x \)-axis can be approximated by the sum of the areas of a set of rectangles as shown by the dashed lines in figure A.8.

Figure A.8: Diagram showing how the area under the curve \( y = f(x) \) may be approximated by a set of rectangles. The intervals \( x_k - x_{k-1} \) need not be the same length.

In general the closed interval \( a \leq x \leq b \) may be partitioned by a set of \( n - 1 \) distinct, ordered points

\[
a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b
\]

to produce \( n \) sub-divisions: in figure A.8 \( n = 6 \) and the spacings are equal. On each interval we construct a rectangle: on the \( k \)th rectangle the height is \( f(l_k) \) chosen to be the smallest value of \( f(x) \) in the interval. These rectangles are shown in the figure.

Another set of rectangles of height \( f(h_k) \) chosen to be the largest value of \( f(x) \) in the interval can also be formed. If \( A \) is the area under the graph it follows that

\[
\sum_{k=1}^{n} (x_k - x_{k-1}) f(l_k) \leq A \leq \sum_{k=1}^{n} (x_k - x_{k-1}) f(h_k).
\]

This type of approximation underlies the simplest numerical methods of approximating integrals and, as will be seen in chapter 4, is the basis of Euler’s approximations to variational problems.
The theory of integration developed by Riemann (1826–1866) shows that for continuous functions these two bounds approach each other, as \( n \to \infty \) in a meaningful manner, and defines the wider class of functions for which this limit exists. When these limits exist their common value is named the integral of \( f(x) \) and is denoted by

\[
\int_a^b dx f(x) \quad \text{or} \quad \int_a^b f(x) \, dx. \tag{A.45}
\]

In this context the function \( f(x) \) is named the integrand, and \( b \) and \( a \) the upper and lower integration limits, or just limits. It can be shown that the integral exists for bounded, piecewise continuous functions and also some unbounded functions.

From this definition the following elementary properties can be derived.

P:I: If \( F(x) \) is a differentiable function and \( F'(x) = f(x) \) then \( F(x) = F(a) + \int_a^x dt \, f(t) \).

This is the Fundamental Theorem of Calculus and is important because it provides one of the most useful tools for evaluating integrals.

P:II: \( \int_a^b dx f(x) = - \int_b^a dx f(x) \).

P:III: \( \int_a^b dx f(x) = \int_a^c dx f(x) + \int_c^b dx f(x) \) provided all integrals exist. Note, it is not necessary that \( c \) lies in the interval \((a, b)\).

P:IV: \( \int_a^b dx (\alpha f(x) + \beta g(x)) = \alpha \int_a^b dx f(x) + \beta \int_a^b dx g(x) \), where \( \alpha \) and \( \beta \) are real or complex numbers.

P:V: \( \left| \int_a^b dx f(x) \right| \leq \int_a^b dx |f(x)| \). This is the analogue of the finite sum inequality

\[
\sum_{k=1}^n a_k \leq \sum_{k=1}^n |a_k|, \quad \text{where } a_k, \, k = 1, 2, \ldots, n, \text{ are a set of complex numbers or functions.}
\]

P:VI: The Cauchy–Schwarz inequality for real functions is

\[
\left( \int_a^b dx f(x)g(x) \right)^2 \leq \left( \int_a^b dx f(x)^2 \right) \left( \int_a^b dx g(x)^2 \right)
\]

with equality if and only if \( g(x) = cf(x) \) for some real constant \( c \). This inequality is sometimes named the Cauchy inequality and sometimes the Schwarz inequality. It is the analogue of the finite sum inequality

\[
\left( \sum_{k=1}^n a_k b_k \right)^2 \leq \left( \sum_{k=1}^n a_k^2 \right) \left( \sum_{k=1}^n b_k^2 \right)
\]

with equality if and only if \( b_k = ca_k \) for all \( k \) and some real constant \( c \).
P:VII: The Hölder inequality: if \( \frac{1}{p} + \frac{1}{q} = 1 \), \( p > 1 \) and \( q > 1 \) then
\[
\int_a^b dx \ |f(x)g(x)| \leq \left( \int_a^b dx \ |f(x)|^p \right)^{1/p} \left( \int_a^b dx \ |g(x)|^q \right)^{1/q},
\]
is valid for complex functions \( f(x) \) and \( g(x) \) with equality if and only if \( |f(x)|^p |g(x)|^{-q} \) and \( \arg(fg) \) are independent of \( x \). It is the analogue of the finite sum inequality
\[
\sum_{k=1}^n |a_k b_k| \leq \left( \sum_{k=1}^n |a_k|^p \right)^{1/p} \left( \sum_{k=1}^n |b_k|^q \right)^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1,
\]
with equality if and only if \( |a_k|^p |b_n|^{-q} \) and \( \arg(a_n b_n) \) are independent of \( n \) (or \( a_k = 0 \) for all \( k \) or \( b_k = 0 \) for all \( k \)). If all \( a_k \) and \( b_k \) are positive and \( p = q = 2 \) these inequalities reduce to the Cauchy–Schwarz inequalities.

P:VIII: The Minkowski inequality for any \( p > 1 \) and real functions \( f(x) \) and \( g(x) \) is
\[
\left( \int_a^b dx \ |f(x) + g(x)|^p \right)^{1/p} \leq \left( \int_a^b dx \ |f(x)|^p \right)^{1/p} + \left( \int_a^b dx \ |g(x)|^p \right)^{1/p},
\]
with equality if and only if \( g(x) = c f(x) \), with \( c \) a non-negative constant. It is the analogue of the finite sum inequality valid for \( a_k, b_k > 0 \), for all \( k \), and \( p > 1 \)
\[
\left( \sum_{k=1}^n (a_k + b_k)^p \right)^{1/p} \leq \left( \sum_{k=1}^n a_k^p \right)^{1/p} + \left( \sum_{k=1}^n b_k^p \right)^{1/p},
\]
with equality if and only if \( b_k = c a_k \) for all \( k \) and \( c \) a non-negative constant.

Sometimes it is convenient to ignore the integration limits, here \( a \) and \( b \), and write \( \int dx f(x) \): this is named the indefinite integral: its value is undefined to within an additive constant. However, it is almost always possible to express problems in terms of definite integrals — that is, those with limits.

The theory of integration is concerned with understanding the nature of the integration process and with extending these simple ideas to deal with wider classes of functions. The sciences are largely concerned with evaluating integrals, that is converting integrals to numbers or functions that can be understood: most of the techniques available for this activity were developed in the nineteenth century or before, and we describe them later in this section.

There are two important extensions to the integral defined above. If either or both \( -a \) and \( b \) tend to infinity we define an infinite integral as a limit of integrals: thus if \( b \to \infty \) we have
\[
\int_a^\infty dx \ f(x) = \lim_{b \to \infty} \left( \int_a^b dx \ f(x) \right), \quad \text{(A.46)}
\]
assuming the limit exists. There are similar definitions for
\[
\int_a^b dx \ f(x) \quad \text{and} \quad \int_{-\infty}^\infty dx \ f(x),
\]
A.2. FUNCTIONS OF A REAL VARIABLE

However, it should be noted that the limit lim_{a \to \infty} \int_{-a}^{a} dx \ f(x) may exist, but the limit lim_{b \to \infty} \int_{-b}^{b} dx \ f(x) may not. An example is \( f(x) = x/(1 + x^2) \) for which

\[
\int_{-b}^{b} dx \ \frac{x}{1 + x^2} = \frac{1}{2} \ln \left( \frac{1 + a^2}{1 + b^2} \right).
\]

If \( a = b \) the right-hand side is zero for all \( a \) (because \( f(x) \) is an odd function) and the first limit is zero: if \( a \neq b \) the second limit does not exist.

Whether or not infinite integrals exist depends upon the behaviour of \( f(x) \) as \( |x| \to \infty \). Consider the limit (A.46). If \( f(x) \neq 0 \) for sufficiently large \( x > 0 \), the limit exists provided \( |f(x)| \to 0 \) faster than \( x^{-\alpha} \), \( \alpha > 1 \): if \( f(x) \) decays to zero slower than \( 1/x^{1-\epsilon} \), for any \( \epsilon > 0 \) the integral diverges, see however exercise A.48, (page 380).

If the integrand is oscillatory, cancellation between the positive and negative parts of the integral gives convergence when the magnitude of the integrand tends to zero. In this case we have the following useful theorem from 1853, due to Chartier\(^7\).

**Theorem A.6** If \( f(x) \to 0 \) monotonically as \( x \to \infty \) and if \( \left| \int_{a}^{x} dt \ \phi(t) \right| \) is bounded as \( x \to \infty \) then \( \int_{a}^{\infty} dx \ f(x) \phi(x) \) exists.

For instance if \( \phi(x) = \sin(\lambda x) \), and \( f(x) = x^{-\alpha} \), \( 0 < \alpha < 2 \) this shows that \( \int_{0}^{\infty} dx \ x^{-\alpha} \sin \lambda x \) exists: if \( \alpha = 1 \) its value is \( \pi/2 \), for any \( \lambda > 0 \). It should be mentioned that the very cancellation which ensures convergence may cause difficulties when evaluating such integrals numerically.

The second important extension deals with integrands that are unbounded. Suppose that \( f(x) \) is unbounded at \( x = a \), then we define

\[
\int_{a}^{b} dx \ f(x) = \lim_{\epsilon \to 0^+} \int_{a+\epsilon}^{b} dx f(x), \quad (A.47)
\]

provided the limit exists. As a general rule, provided \( |f(x)| \) tends to infinity slower than \( |x - a|^\beta \), \( \beta > -1 \), the integral exists, which is why, in the previous example, we needed \( \alpha < 2 \); note that if \( f(x) = O(\ln(x-a)) \), as \( x \to a \), it is integrable. For functions unbounded at an interior point the natural extension to P:III is used.

The evaluation of integrals of any complexity in closed form is normally difficult, or impossible, but there are a few tools that help. The main technique is to use the Fundamental Theorem of Calculus in reverse and simply involves recognising those \( F(x) \) whose derivative is the integrand: this requires practice and ingenuity. The main purpose of the other tools is to convert integrals into recognisable types. The first is *integration by parts*, derived from the product rule for differentiation:

\[
\int_{a}^{b} dx \ u \ \frac{dv}{dx} = [uv]_{a}^{b} - \int_{a}^{b} dx \ \frac{du}{dx} v. \quad (A.48)
\]

The second method is to change variables:

\[
\int_{a}^{b} dx \ f(x) = \int_{A}^{B} dt \ \frac{dx}{dt} f(g(t)) = \int_{A}^{B} dt \ g'(t)f(g(t)). \quad (A.49)
\]

\(^7\)Chartier, Journal de Math. 1853, XVIII, pages 201–212.
where \( x = g(t), g(A) = a, g(B) = b \), and \( g(t) \) is monotonic for \( A < t < B \). In these circumstances the Leibniz notation is helpfully transparent because \( \frac{dx}{dt} \) can be treated like a fraction, making the equation easier to remember. The geometric significance of this formula is simply that the small element of length \( \delta x \) at \( x \), becomes the element of length \( \delta x = g'(t)\delta t \), where \( x = g(t) \), under the variable change.

The third method involves the differentiation of a parameter. Consider a function \( f(x, u) \) of two variables, which is integrated with respect to \( x \), then

\[
\frac{d}{du} \int_{a(u)}^{b(u)} dx f(x, u) = f(b, u) \frac{db}{du} - f(a, u) \frac{da}{du} + \int_{a(u)}^{b(u)} dx \frac{\partial f}{\partial u},
\]

(A.50)

provided \( a(u) \) and \( b(u) \) are continuously differentiable and \( f_u(x, u) \) is a continuous function of both variables; the derivation of this formula is considered in exercise A.46. If neither limit depends upon \( u \) the first two terms on the right-hand side vanish. A simple example shows how this method can work. Consider the integral

\[
I(u) = \int_{0}^{\infty} dx \; e^{-x^u}, \quad u > 0.
\]

The derivatives are

\[
I'(u) = - \int_{0}^{\infty} dx \; xe^{-x^u} \quad \text{and, in general,} \quad I^{(n)}(u) = (-1)^n \int_{0}^{\infty} dx \; x^n e^{-x^u}.
\]

But the original integral is trivially integrated to \( I(u) = 1/u \), so differentiation gives

\[
\int_{0}^{\infty} dx \; x^n e^{-x^u} = \frac{n!}{u^{n+1}}.
\]

This result may also be found by repeated integration by parts but the above method involves less algebra. The application of these methods usually requires some skill, some trial and error and much patience. Please do not spend too long on the following problems.

**Exercise A.38 (B)**

(a) If \( f(x) \) is an odd function, \( f(-x) = -f(x) \), show that \( \int_{-a}^{a} dx \; f(x) = 0 \).

(b) If \( f(x) \) is an even function, \( f(-x) = f(x) \), show that \( \int_{-a}^{a} dx \; f(x) = 2 \int_{0}^{a} dx \; f(x) \).\( \square \)

**Exercise A.39 (B)** Show that, if \( \lambda > 0 \), the value of the integral \( I(\lambda) = \int_{0}^{\infty} dx \; \frac{\sin \lambda x}{x} \) is independent of \( \lambda \). How are the values of \( I(\lambda) \) and \( I(-\lambda) \) related? \( \square \)

**Exercise A.40 (B,E)** Use integration by parts to evaluate the following indefinite integrals.

(a) \( \int dx \; \ln x \), \quad (b) \( \int dx \; \frac{x}{\cos^2 x} \), \quad (c) \( \int dx \; x \ln x \), \quad (d) \( \int dx \; x \sin x \). \( \square \)

**Exercise A.41 (B)** Evaluate the following integrals

(a) \( \int_{0}^{\pi/4} dx \; \sin x \ln(\cos x) \), \quad (b) \( \int_{0}^{\pi/4} dx \; x \tan^2 x \), \quad (c) \( \int_{0}^{1} dx \; x^2 \sin^{-1} x \). \( \square \)
Exercise A.42 (B,E) If \( I_n = \int_0^x t^n e^{at} \, dt \), \( n \geq 0 \), use integration by parts to show that \( aI_n = x^n e^{ax} - nI_{n-1} \) and deduce that
\[
I_n = n! e^{ax} \sum_{k=0}^{n} \frac{(-1)^{n-k}}{a^{n-k+1} k!} x^k - \frac{(-1)^n n!}{a^{n+1}}.
\]

Exercise A.43 (B)
(a) Using the substitution \( u = a - x \), show that \( \int_0^a dx f(x) = \int_0^a dx f(a - x) \).

(b) With the substitution \( \theta = \pi/2 - \phi \) show that
\[
I = \int_0^{\pi/2} d\theta \frac{\sin \theta}{\sin \theta + \cos \theta} = \int_0^{\pi/2} d\phi \frac{\cos \phi}{\cos \phi + \sin \phi}
\]
and deduce that \( I = \pi/4 \).

Exercise A.44 (B) Use the substitution \( t = \tan(x/2) \) to prove that if \( a > |b| > 0 \)
\[
\int_0^\pi dx \frac{1}{a + b \cos x} = \frac{\pi}{\sqrt{a^2 - b^2}}.
\]
Why is the condition \( a > |b| \) necessary?

Use this result and the technique of differentiating the integral to determine the values of,
\[
\int_0^\pi \frac{dx}{(a + b \cos x)^2}, \quad \int_0^\pi \frac{dx}{(a + b \cos x)^3}, \quad \int_0^\pi \frac{dx}{(a + b \cos x)^2}, \quad \int_0^\pi dx \ln(a + b \cos x).
\]

Exercise A.45 (B) Prove that \( y(t) = \frac{1}{\omega} \int_a^t dx f(x) \sin \omega(t - x) \) is the solution of the differential equation
\[
d^2y/dt^2 + \omega^2 y = f(t), \quad y(a) = 0, \quad y'(a) = 0.
\]

Exercise A.46 (B) (a) Consider the integral \( F(u) = \int_0^{a(u)} dx f(x) \), where only the upper limit depends upon \( u \). Using the basic definition, equation (A.5) (page 354), derive the derivative \( F'(u) \).

(b) Consider the integral \( F(u) = \int_a^b dx f(x, u) \), where only the integrand depends upon \( u \). Using the basic definition derive the derivative \( F'(u) \).

Exercise A.47 (B) Assuming that both integrals exist, show that
\[
\int_{-\infty}^{\infty} dx f \left( x - \frac{1}{x} \right) = \int_{-\infty}^{\infty} dx f(x).
\]
Hence show that
\[
\int_{-\infty}^{\infty} dx \exp \left( -x^2 - \frac{1}{x^2} \right) = \frac{\sqrt{\pi}}{e^2}.
\]
You will need the result \( \int_{-\infty}^{\infty} dx \ e^{-x^2} = \sqrt{\pi} \).
Exercise A.48 (B) Find the limits as $X \to \infty$ of the following integrals

$$\int_2^X \frac{1}{x \ln x} \, dx \quad \text{and} \quad \int_2^X \frac{1}{x (\ln x)^2} \, dx.$$

Hint note that if $f(x) = \ln(\ln x)$ then $f'(x) = (x \ln x)^{-1}$.

Exercise A.49 (B) Determine the values of the real constants $a > 0$ and $b > 0$ for which the following limit exists

$$\lim_{X \to \infty} \int_2^X \frac{1}{x^a (\ln x)^b}.$$

A.3 Miscellaneous exercises

The following exercises can be tackled using the method described in the corresponding section, though other methods may also be applicable.

Limits

Exercise A.50 (B) Find, using first principles, the following limits

(a) $\lim_{x \to 1} \frac{x^a - 1}{x - 1}$,  \hspace{1cm} (b) $\lim_{x \to 0} \frac{\sqrt{1 + x} - 1}{\sqrt{x}}$,  \hspace{1cm} (c) $\lim_{x \to a} \frac{x^{1/3} - a^{1/3}}{x^{1/2} - a^{1/2}}$,

(d) $\lim_{x \to (\pi/2)^-} (\pi - 2x) \tan x$,  \hspace{1cm} (e) $\lim_{x \to 0_+} x^{1/x}$,  \hspace{1cm} (f) $\lim_{x \to b} \left( \frac{1 + x}{1 - x} \right)^{1/x}$

where $a$ is a real number.

Inverse functions

Exercise A.51 (B) Show that the inverse functions of $y = \cosh x$, $y = \sinh x$ and $y = \tanh x$, for $x > 0$ are, respectively

$$x = \ln \left( y + \sqrt{y^2 - 1} \right), \quad x = \ln \left( y + \sqrt{y^2 + 1} \right) \quad \text{and} \quad x = \frac{1}{2} \ln \left( \frac{1 + y}{1 - y} \right).$$

Exercise A.52 (B) The function $y = \sin x$ may be defined to be the solution of the differential equation

$$\frac{d^2 y}{dx^2} + y = 0, \quad y(0) = 0, \quad y'(0) = 1.$$

Show that the inverse function $x(y)$ satisfies the differential equation

$$\frac{d^2 x}{dy^2} = y \left( \frac{dx}{dy} \right)^3$$

which gives $x(y) = \sin^{-1} y = \int_0^y \frac{du}{\sqrt{1 - u^2}}$.

Hence find the Taylor series of $\sin^{-1} y$ to $O(y^5)$.

Hint you may find it helpful to solve the equation by defining $z = dx/dy$. 

\[\square\]
Derivatives

Exercise A.53 (B) Find the derivative of \( y(x) \) where
(a) \( y = f(x)^{g(x)} \), (b) \( y = \sqrt{\frac{p+x}{p-x}} \sqrt{\frac{q+x}{q-x}} \), (c) \( y^n = x + \sqrt{1 + x^2} \).
\[ \square \]

Exercise A.54 (B) If \( y = \sin(a \sin^{-1} x) \) show that \((1 - x^2)y'' - xy' + a^2y = 0\).
\[ \square \]

Exercise A.55 (B) If \( g(x) \) satisfies the equation \((1 - x^2)\frac{d^2y}{dx^2} - 2x \frac{dy}{dx} + \lambda y = 0\), where \( \lambda \) is a constant and \(|x| \leq 1\), show that changing the independent variable, \( x \), to \( \theta \) where \( x = \cos \theta \) changes this to
\[ \frac{d^2y}{d\theta^2} + \cot \theta \frac{dy}{d\theta} + \lambda y = 0. \]
\[ \square \]

Exercise A.56 (B) The Schwarzian derivative of a function \( f(x) \) is defined to be
\[ Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2} \left( \frac{f''(x)}{f'(x)} \right)^2 = -2\sqrt{f'(x)} \frac{d^2}{dx^2} \left( \frac{1}{\sqrt{f'(x)}} \right). \]
Show that if \( f(x) \) and \( g(x) \) both have negative Schwarzian derivatives, \( Sf(x) < 0 \) and \( Sg(x) < 0 \), then the Schwarzian derivative of the composite function \( h(x) = f(g(x)) \) also satisfies \( Sh(x) < 0 \).
Note the Schwarzian derivative is important in the study of the fixed points of maps.
\[ \square \]

Partial derivatives

Exercise A.57 (B,E) If \( z = f(x + ay) + g(x - ay) - \frac{x}{2a^2} \cos(x + ay) \) where \( f(u) \) and \( g(u) \) are arbitrary functions of a single variable and \( a \) is a constant, prove that
\[ a^2 \frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} = \sin(x + ay). \]
\[ \square \]

Exercise A.58 (B) If \( f(x, y, z) = \exp(ax + by + cz)/xyz \), where \( a, b \) and \( c \) are constants, find the partial derivatives \( f_x, f_y \) and \( f_z \), and solve the equations \( f_x = 0, f_y = 0 \) and \( f_z = 0 \) for \((x, y, z)\).
\[ \square \]

Exercise A.59 (B,E) The equation \( f(u^2 - x^2, u^2 - y^2, u^2 - z^2) = 0 \) defines \( u \) as a function of \( x, y \) and \( z \).
Show that \[ \frac{1}{x} \frac{\partial u}{\partial x} + \frac{1}{y} \frac{\partial u}{\partial y} + \frac{1}{z} \frac{\partial u}{\partial z} = \frac{1}{u}. \]
\[ \square \]

Implicit functions

Exercise A.60 (B) Show that the function \( f(x, y) = x^2 + y^2 - 1 \) satisfies the conditions of the Implicit Function Theorem for most values of \((x, y)\), and that the function \( y(x) \) obtained from the theorem has derivative \( y'(x) = -x/y \).
The equation \( f(x, y) = 0 \) can be solved explicitly to give the equations \( y = \pm \sqrt{1 - x^2} \). Verify that the derivatives of both these functions is the same as that obtained from the Implicit Function Theorem.
\[ \square \]
Exercise A.61 (B) Prove that the equation \( x \cos xy = 0 \) has a unique solution, \( y(x) \), near the point \( (1, \frac{x}{2}) \), and find its first and second derivatives.

Exercise A.62 (B,E) The folium of Descartes has equation \( f(x, y) = x^3 + y^3 - 3axy = 0 \). Show that at all points on the curve where \( y^2 \neq ax \), the implicit function \( y(x) \) has derivative
\[
\frac{dy}{dx} = \frac{-x^2 - ay}{y^2 - ax}.
\]
Show that there is a horizontal tangent to the curve at \( (a^{2^{1/3}}, a^{4^{1/3}}) \).

Taylor series

Exercise A.63 (B) By sketching the graphs of \( y = \tan x \) and \( y = \frac{1}{x} \) for \( x > 0 \) show that the equation \( x \tan x = 1 \) has an infinite number of positive roots. By putting \( x = n\pi + z \), where \( n \) is a positive integer, show that this equation becomes \( (n\pi + z)\tan z = 1 \) and use a first order Taylor expansion of this to show that the root nearest \( n\pi \) is given approximately by \( x_n = n\pi + \frac{1}{n\pi} \).

Exercise A.64 (B) Determine the constants \( a \) and \( b \) such that \( \frac{1 + a \cos 2x + b \cos 4x}{x^4} \) is finite at the origin.

Exercise A.65 (B) Find the Taylor series, to 4th order, of the following functions:
(a) \( \ln \cosh x \), (b) \( \ln(1 + \sin x) \), (c) \( e^x \sin x \), (d) \( \sin^2 x \).

Mean value theorem

Exercise A.66 (B) If \( f(x) \) is a function such that \( f'(x) \) increases with increasing \( x \), use the Mean Value theorem to show that \( f'(x) < f(x + 1) - f(x) < f'(x + 1) \).

Exercise A.67 (B) Use the functions \( f_1(x) = \ln(1 + x) - x \) and \( f_2(x) = f_1(x) + x^2/2 \) and the Mean Value Theorem to show that, for \( x > 0 \),
\[
x - \frac{1}{2}x^2 < \ln(1 + x) < x.
\]

L’Hospital’s rule

Exercise A.68 (B) Show that \( \lim_{x \to 1} \frac{\sin \ln x}{x^5 - 7x^3 + 6} = -\frac{1}{16} \).

Exercise A.69 (B) Determine the limits \( \lim_{x \to 0} (\cos x)^{1/\tan^2 x} \) and \( \lim_{x \to 0} \frac{a \sin bx - b \sin ax}{x^3} \).

Integrals

Exercise A.70 (B) Using differentiation under the integral sign show that
\[
\int_0^\infty dx \frac{\tan^{-1}(ax)}{x(1 + x^2)} = \frac{1}{2} \pi \ln(1 + a).
\]
Exercise A.71 (B) Prove that, if $|a| < 1$

$$\int_{0}^{\pi/2} dx \frac{\ln (1 + \cos \pi a \cos x)}{\cos x} = \frac{\pi^2}{8} (1 - 4a^2).$$

Exercise A.72 (B) If $f(x) = (\sin x)/x$, show that

$$\int_{0}^{\pi/2} dx f(x) f(\pi/2-x) = \frac{2}{\pi} \int_{0}^{\pi} dx f(x).$$

Exercise A.73 (B) Use the integral definition

$$\tan^{-1} x = \int_{0}^{x} dt \frac{1}{1 + t^2}$$

to show that, for $x > 0$,

$$\tan^{-1}(1/x) = \int_{x}^{\infty} dt \frac{1}{1 + t^2}$$

and deduce that $\tan^{-1} x + \tan^{-1}(1/x) = \pi/2$.

Exercise A.74 (B) Determine the values of $x$ that make $g'(x) = 0$ if $g(x) = \int_{x}^{2x} dt f(t)$ and (a) $f(t) = e^t$, and (b) $f(t) = (\sin t)/t$.

Exercise A.75 (B) If $f(x)$ is integrable for $a \leq x \leq a + h$ show that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f \left( a + \frac{kh}{n} \right) = \frac{1}{h} \int_{a}^{a+h} dx f(x).$$

Hence find the following limits

(a) $\lim_{n \to \infty} n^{-6} \left( 1 + 2^5 + 3^5 + \cdots + n^5 \right)$, (b) $\lim_{n \to \infty} \left( \frac{1}{1 + n} + \frac{1}{2 + n} + \cdots + \frac{1}{3n} \right)$,

(c) $\lim_{n \to \infty} \frac{1}{n} \left( \sin \frac{y}{n} + \sin \frac{2y}{n} + \cdots + \sin y \right)$, (d) $\lim_{n \to \infty} n^{-1} [(n + 1)(n + 2) \cdots (2n)]^{1/n}$

Exercise A.76 (B) If the functions $f(x)$ and $g(x)$ are differentiable find expressions for the first derivative of the functions

$$F(u) = \int_{0}^{u} dx \frac{f(x)}{\sqrt{u^2 - x^2}} \quad \text{and} \quad G(u) = \int_{0}^{u} dx \frac{g(x)}{(u - x)^a} \quad \text{where} \quad 0 < a < 1.$$

This is a fairly difficult problem. The formula (A.50) does not work because the integrands are singular, yet by substituting simple functions for $f(x)$ and $g(x)$, for instance $1$, $x$ and $x^2$, we see that there are cases for which the functions $F(u)$ and $G(u)$ are differentiable. Thus we expect an equivalent to formula (A.50) to exist.