Contents

1 Preliminary Analysis **5**

1.1 Introduction 5
1.2 Notation and preliminary remarks 8
1.2.1 The Order notation 10
1.3 Functions of a real variable 12
1.3.1 Introduction 12
1.3.2 Continuity and Limits 12
1.3.3 Monotonic functions and inverse functions 14
1.3.4 The derivative 15
1.3.5 Mean Value Theorems 19
1.3.6 Partial Derivatives 21
1.3.7 Implicit functions 25
1.3.8 Taylor series for one variable 28
1.3.9 Taylor series for several variables 33
1.3.10 L'Hospital's rule 34
1.3.11 Integration 35
1.4 Miscellaneous exercises 42
1.5 Solutions for chapter 1 47

2 The Calculus of Variations **73**

2.1 Introduction 73
2.2 The shortest distance between two points in a plane 73
2.2.1 The stationary distance 74
2.2.2 The shortest path: local and global minima 76
2.2.3 Gravitational Lensing 78
2.3 Two generalisations 79
2.3.1 Functionals depending only upon $y'(x)$ 79
2.3.2 Functionals depending upon x and $y'(x)$ 81
2.4 Notation 82
2.5 Examples of functionals 83
2.5.1 The brachistochrone 84
2.5.2 Minimal surface of revolution 85
2.5.3 The minimum resistance problem 86
2.5.4 A problem in navigation 90
2.5.5 The isoperimetric problem 90

2.5.6 The catenary . . . 91
2.5.7 Fermat's principle . . . 92
2.5.8 Coordinate free formulation of Newton's equations . . . 93
2.6 Miscellaneous exercises . . . 95
2.7 Solutions for chapter 2 . . . 99

3 The Euler-Lagrange equation 109
3.1 Introduction . . . 109
3.2 Preliminary remarks . . . 110
3.2.1 Relation to differential calculus . . . 110
3.2.2 Differentiation of a functional . . . 111
3.3 The fundamental lemma . . . 115
3.4 The Euler-Lagrange equations . . . 116
3.4.1 The first integral . . . 118
3.5 Theorems of Bernstein and du Bois-Reymond . . . 121
3.5.1 Bernstein's theorem . . . 122
3.5.2 The contrast between initial and boundary value problems . . . 123
3.6 Strong and Weak variations . . . 125
3.7 Miscellaneous exercises . . . 128
3.8 Solutions for chapter 3 . . . 134

4 Applications of the Euler-Lagrange equation 151
4.1 Introduction . . . 151
4.2 The brachistochrone . . . 151
4.2.1 The cycloid . . . 152
4.2.2 Formulation of the problem . . . 155
4.2.3 A solution . . . 156
4.3 Minimal surface of revolution . . . 160
4.3.1 Derivation of the functional . . . 160
4.3.2 Applications . . . 162
4.3.3 The solution in a special case . . . 162
4.3.4 Summary . . . 166
4.4 Soap Films . . . 169
4.5 Miscellaneous exercises . . . 173
4.6 Solutions for chapter 4 . . . 176

5 Further theoretical developments 193
5.1 Introduction . . . 193
5.2 Invariance of the Euler-Lagrange equation . . . 194
5.2.1 Changing the independent variable . . . 194
5.2.2 Changing both the dependent and independent variables . . . 197
5.3 Functionals containing many dependent variables . . . 202
5.3.1 Introduction . . . 202
5.3.2 Two dependent variables . . . 203
5.3.3 Many dependent variables . . . 206
5.3.4 Changing dependent variables . . . 208
5.4 Symmetries . . . 209
5.4.1 Invariance under translations . . . 210

5.4.2 Noether's theorem . . . 212
5.4.3 Proof of Noether's theorem . . . 218
5.5 The Inverse Problem . . . 220
5.6 Miscellaneous exercises . . . 222
5.7 Solutions for chapter 5 . . . 226

6 The second variation 245
6.1 Introduction . . . 245
6.2 Stationary points of functions of several variables . . . 246
6.2.1 Functions of one variable . . . 246
6.2.2 Functions of two variables . . . 247
6.2.3 Functions of n variables . . . 248
6.3 The second variation of a functional . . . 251
6.3.1 Short intervals . . . 252
6.3.2 Legendre's necessary condition . . . 254
6.4 Analysis of the second variation . . . 255
6.4.1 Analysis of Δ_2 . . . 257
6.5 The Variational Equation . . . 261
6.6 The Brachistochrone problem . . . 262
6.7 Surface of revolution . . . 264
6.8 The connection between Jacobi's equation and quadratic forms . . . 266
6.9 Appendix: Riccati's equation . . . 268
6.10 Miscellaneous exercises . . . 269
6.11 Solutions for chapter 6 . . . 272

7 Parametric Functionals 283
7.1 Introduction: parametric equations . . . 283
7.2 The parametric variational problem . . . 287
7.2.1 Geodesics . . . 290
7.2.2 The Brachistochrone problem . . . 292
7.2.3 Surface of Minimum Revolution . . . 293
7.3 The parametric and the conventional formulation . . . 293
7.4 Miscellaneous exercises . . . 295
7.5 Solutions for chapter 7 . . . 299

8 Free boundaries etc 309
8.1 Introduction . . . 309

9 Differential Calculus in Linear Spaces 311
9.1 Introduction . . . 311

10 Conditional Stationary problems 313
10.1 Introduction . . . 313
10.2 Lagrange's undetermined multipliers . . . 317
10.3 End of chapter exercises . . . 323

11 Sturm-Liouville systems **327**
11.1 Introduction 327
11.2 Conditional Stationary values of functionals 327
11.2.1 An isoperimetric problem 331
11.3 Sturm-Liouville systems: eigenvalues and eigenfunctions 335
11.3.1 Introduction 335
11.3.2 Definition of a Sturm-Liouville system 338
11.3.3 Fourier series 345
11.3.4 Special functions 349
11.4 Appendix 351
11.5 End of chapter exercises 355

12 Direct Methods **357**
12.0.1 Introduction 357
12.0.2 Direct approximation for the eigenvalues of Sturm-Liouville problems 363
12.0.3 General boundary conditions for a Sturm-Liouville system 367

Chapter 1

Preliminary Analysis

1.1 Introduction

This course is about two related mathematical concepts which are of use in many areas of applied mathematics, are of immense importance in formulating the laws of theoretical physics and also produce important, interesting and some unsolved mathematical problems. These are the *functional* and *variational principles*: the theory of these entities is named *The Calculus of Variations*.

A functional is a generalisation of a function of one or more real variables. A real function of a single real variable maps an interval of the real line to real numbers: for instance, the function $(1+x^2)^{-1}$ maps the whole real line to the interval $(0,1]$; the function $\ln x$ maps the positive real axis to the whole real line. Similarly a real function of n real variables maps a domain of R^n into the real numbers.

A *functional* maps a given class of functions to real numbers. A simple example of a functional is

label: eq:vp1-intr01

$$S[y] = \int_0^1 dx\, \sqrt{1+y'(x)^2}, \quad y(0)=0, \quad y(1)=1, \tag{1.1}$$

which associates a real number with any real function $y(x)$ which satisfies the boundary conditions and for which the integral exists. We use the square bracket notation[1] $S[y]$ to emphasise the fact that the functional depends upon the choice of function used to evaluate the integral. In chapter 2 we shall see that a wide variety of problems can be described in terms of functionals. Notice that the boundary conditions, $y(0)=0$ and $y(1)=1$ in this example, are normally part of the definition of the functional.

Real functions of n real variables can have various properties; for instance they can be continuous, they may be differentiable or they may have stationary points and local and global maxima and minima: functionals share many of these properties. In

[1] In this course we use conventions common in applied mathematics and theoretical physics. A function of a real variable x will usually be represented by symbols such as $f(x)$ or just f, there often being no distinction made between the function and its value; as is often the case it is often clearer to use context to provide meaning, rather than precise definitions, which initially can hinder clarity. Similarly, we use the older convention, $S[y]$, for a functional, to emphasise that y is itself a function; this distinction is not made in modern mathematics. For an introductory course we feel that the older convention, used in most texts, is clearer and more helpful.

particular the notion of a stationary point of a function has an important analogy in the theory of functionals and this gives rise to the idea of a *variational principle*, which arises when the solution to a problem is given by the function making a particular functional stationary. Variational principles are common and important in the natural sciences.

The simplest example of a variational principle is that of finding the shortest distance between two points. Suppose the two points lie in a plane, with one point at the origin, O, and the other at point A with coordinates $(1, 1)$, then if $y(x)$ represents a smooth curve passing through O and A the distance between O and A, along this curve is given by the functional defined in equation 1.1. The shortest path is that which minimises the value of $S[y]$. If the surface is curved, for instance a sphere or ellipsoid, the equivalent functional is more complicated, but the shortest path is that which minimises it.

Variational principles are important for three principal reasons. First, many problems are naturally formulated in terms of a functional and an associated variational principle. Several of these will be described in chapter 2 and some solutions will be obtained as the course develops.

Second, most equations of mathematical physics can be derived from variational principles. This is important partly because it suggests a unifying theme in our description of nature and partly because such formulations are independent of any particular coordinate system, so making the essential mathematical structure of the equations more transparent and easier to understand. This aspect of the subject is not considered in this course; a good discussion of these problems can be found in Yourgrau and Mandelstam (1968)[2].

Finally, variational principles provide powerful computational tools; we explore aspects of this theory in chapter 12.

Consider the problem of finding the shortest path between two points on a curved surface. The associated functional assigns a real number to each smooth curve joining the points. A first step to solving this problem is to find the stationary values of the functional; it is then necessary to decide which of these provides the shortest path. This is very similar to the problem of finding extreme values of a function of n variables, where we first determine the stationary points and then classify them: the important and significant difference is that the space of allowed functions is not usually finite in dimension. The infinite dimensional spaces of functions, with which we shall be dealing, has many properties similar to those possessed by finite dimensional spaces, and in the many problems the difference is not significant. However, this generalisation does introduce some practical and technical difficulties, which we shall consider later in the course. In this chapter we review calculus in order to prepare for these more general ideas of calculus.

In elementary calculus and analysis, the functions studied first are 'real functions, f, of one real variable', that is, functions with domain either R or a subset of R, and codomain R. Without any other restrictions on f, this definition is too general to be useful in calculus and applied mathematics. Most functions of one real variable that are of interest in applications have smooth graphs, although sometimes they may fail to be smooth at one or more points where they have a 'kink' (fail to be differentiable), or even a break (where they are discontinuous). This smooth behaviour is related to

[2] Yourgrau W and Mandelstram S 1968 *Variational Principles in Dynamics and Quantum Theory* (Pitman), reprinted by Dover 1979.

the fact that most important functions of one variable describe physical phenomena and often arise as solutions of ordinary differential equations. Therefore it is usual to restrict attention to functions that are differentiable or, more usually, differentiable a number of times.

The most useful generalisation of differentiability to functions defined on sets other than R requires some care. It is not too hard in the case of functions of several (real) variables but we shall have to generalise differentiation and integration to *functionals*, not just to functions of several real variables.

Our presentation conceals very significant intellectual achievements made at the end of the nineteenth century and during the first half of the twentieth century. During the nineteenth century, although much work was done on particular equations, there was little systematic theory. This changed when the idea of infinite dimensional vector spaces began to emerge. Between 1900 and 1906, fundamental papers appeared by Fredholm[3], Hilbert[4], and Fréchet[5]. Fréchet's thesis gave for the first time definitions of limit and continuity that were applicable in very general sets. Previously, the concepts had been restricted to special objects such as points, curves, surfaces or functions. By introducing the concept of distance in more general sets he paved the way for rapid advances in the theory of partial differential equations. We shall study *normed vector spaces* in chapter 9, which are special cases of what are now called Fréchet spaces. These ideas together with the theory of Lebesgue integration introduced, in 1902, by Lebesgue in his doctoral thesis[6], led to the modern theory of functional analysis. This is now the usual framework of the theoretical study of partial differential equations. They are required also for an elucidation of some of the difficulties in the Calculus of Variations. However, in this introductory course, we concentrate on basic techniques of solving practical problems, because we think this is the best way to motivate and encourage further study.

This preliminary chapter, which is assessed, is about real analysis and introduces many of the ideas needed for our treatment of the Calculus of Variations. It is possible that you are already familiar with the mathematics described in this chapter, in which case you could start the course with chapter 2. You should ensure, however, that you have a good working knowledge of differentiation, both ordinary and partial, Taylor series of one and several variables and differentiation under the integral sign, all of which are necessary for the development of the theory.

Very many exercises are set, in the belief that mathematical ideas cannot be understood without attempting to solve problems at various levels of difficulty and that one learns most by making one's own mistakes, which is time consuming. You should not attempt all these exercise at a first reading, but these provide practice of essential mathematical techniques and in the use of a variety of ideas, so you should do as many as time permits; thinking about a problem, then looking up the solution is usually of

[3] I. Fredholm, *On a new method for the solution of Dirichlet's problem*, reprinted in Oeuvres Complètes, l'Institut Mittag-Leffler, (Malmö) 1955, pp 61-68 and 81-106

[4] D. Hilbert published six papers between 1904 and 1906. They were republished as *Grundzüge einer allgemeinen Theorie der Integralgleichungen* by Teubner, (Leipzig and Berlin), 1924. The most crucial paper is the fourth.

[5] M. Fréchet, Doctoral thesis, *Sur quelques points du Calcul fonctionnel*, Rend. Circ. mat. Palermo 22 (1906), pp 1-74.

[6] H. Lebesgue, Doctoral thesis, Paris 1902, reprinted in Annali Mat. pura e appl., 7 (1902) pp 231-359.

little value. The exercises at the end of this chapter are examples of the type of problem that commonly occur in applications: they are provided for extra practice if time permits and it is *not* necessary for you to attempt them.

1.2 Notation and preliminary remarks

label: sec:vp1-nota

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

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 $f : X \to Y$, 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 recent years this has come to be regarded as heresy by some: however, there are good practical reasons for using this freer notation that do not affect pure mathematics. 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 1.5 (page 25).

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 R^n are represented by n-tuples of real numbers $(x_1, x_2, \dots, x_n)$. It is convenient to use bold faced symbols, $\mathbf{x}$, $\mathbf{a}$ and so on, to denote these points, so $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and we shall write $\mathbf{x}$ and $(x_1, x_2, \dots, x_n)$ interchangeably. In hand-written text a bold character, $\mathbf{x}$, is usually denoted by an underline, $\underline{x}$.

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

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 $a \le x \le b$. By convention, the intervals $(-\infty, a), (b, \infty)$ and $(-\infty,\infty) = R$ are also open intervals. Here, $(-\infty, a)$ means the set of all real numbers strictly less than a. The symbol ∞ for 'infinity', is not a number, and its use here is conventional. In the language and notation of set theory, we can write $(-\infty, a) = \{x \in R : x < a\}$, 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 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 R^n, and for every real α, $\mathbf{x}+\mathbf{y}$ and $\alpha\mathbf{x}$ are in R^n. Explicitly,

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

and

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

Functions $\mathbf{f} : R^n \to R^m$ may also be added and multiplied by real numbers. Therefore a functions of this type may be regarded as a vector in the vector space of functions — though this space is not finite dimensional like R^n. The axiomatic definition of a linear vector space is provided in chapter 9.

In the space 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 label: eq:vp1-edist

$$|\mathbf{x}-\mathbf{y}| = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2}. \qquad (1.2)$$

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) $|\mathbf{x}| \ge 0$, and $|\mathbf{x}| = 0$ if and only if $\mathbf{x} = 0$,
(b) $|\mathbf{x}-\mathbf{y}| = |\mathbf{y}-\mathbf{x}|$,
(c) $|\mathbf{x}-\mathbf{y}| + |\mathbf{y}-\mathbf{z}| \ge |\mathbf{x}-\mathbf{z}|$, (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 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 3 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*. label: def:vp1-nota02

Definition 1.1
The **open ball** $B_r(\mathbf{a})$ of radius r and centre $\mathbf{a} \in R^n$ is the set of points

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

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*

$$\{(x, y) \in R : |x| < 1, |y| < 1\}$$

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*[7], which we can now define.

label: def:vp1-nota03

Definition 1.2
Open sets A set U in R^n is said to be *open* if for every $\mathbf{x} \in U$ there is an open ball $B_r(\mathbf{a})$ wholly contained within U which contains $\mathbf{x}$.

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[8]. Open balls are not the only type of open sets and is not hard to show that the open square, $\{(x, y) \in R : |x| < 1, |y| < 1\}$, 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.

label: ex:vp1-nota01

Exercise 1.1
Show that the open square is an open set by constructing explicitly for each (x, y) in the open square $\{(x, y) \in R : |x| < 1, |y| < 1\}$ a ball containing (x, y) and lying in the square.

1.2.1 The Order notation

It is often useful to have a bound for the magnitude of a function that does not require exact calculation. For example, the function $f(x) = \sqrt{\sin(x^2 \cosh x) - x^2 \cos x}$ tends to zero at a similar rate to x^2 as $x \to 0$ and this information is sometimes more helpful than the detailed knowledge of the function. The *order notation* is designed for this purpose.

Definition 1.3
Order notation. A function $f(x)$ is said to be *of order* x^n as $x \to 0$ if there is a non-zero constant C such that $|f(x)| < C|x^n|$ for all x in an interval around $x = 0$. This is written as

label: eq:vp1-ord01

$$f(x) = O(x^n) \quad \text{as} \quad x \to 0. \tag{1.3}$$

The conditional clause 'as $x \to 0$' is often omitted when it is clear from the context. More generally, this order notation can be used to compare the size of functions, $f(x)$

[7] As with many other concepts in analysis, formulating clearly the concepts, in this case an open set, represents a major achievement.

[8] See for example W A Sutherland, *Introduction to Metric and Topological Spaces*, Oxford University Press.

and $g(x)$: we say that $f(x)$ is of the order of $g(x)$ as $x \to y$ if there is a non-zero constant C such that $|f(x)| < C|g(x)|$ for all x in an interval around y; more succinctly, $f(x) = O(g(x))$ as $x \to y$.

When used in the form $f(x) = O(g(x))$ as $x \to \infty$, this notation means that $|f(x)| < C|g(x)|$ for all $x > X$, where X and C are positive numbers independent of x.

This notation is particularly useful when truncating power series: thus, the series for $\sin x$ up to $O(x^3)$ is written,

$$\sin x = x - \frac{x^3}{3!} + O(x^5),$$

meaning that the remainder is smaller than $C|x|^5$, as $x \to 0$ for some C. Note that this *includes* the x^3 term.

label: ex:vp1-ord01

Exercise 1.2
Show that if $f(x) = O(x^2)$ as $x \to 0$ then also $f(x) = O(x)$.

label: ex:vp1-ord02

Exercise 1.3
Determine the order of the following expressions as $x \to 0$.

$$\text{(a)}\quad x\sqrt{1+x^2}, \quad \text{(b)}\quad \frac{x}{1+x}, \quad \text{(c)}\quad \frac{x^{3/2}}{1-e^{-x}}.$$

label: ex:vp1-ord03

Exercise 1.4
Determine the order of the following expressions as $x \to \infty$.

$$\text{(a)}\quad \frac{x}{x-1}, \quad \text{(b)}\quad \sqrt{4x^2+x} - 2x, \quad \text{(c)}\quad (x+b)^a - x^a, \quad a > 0.$$

The order notation is usefully extended to functions of n real variables, $f : R^n \to R$, by using the distance $|\mathbf{x}|$. Thus, we say that $f(\mathbf{x}) = O(|\mathbf{x}|^n)$ if there is a non-zero constant C and a small number δ such that $|f(\mathbf{x})| < C|\mathbf{x}|^n$ for $|\mathbf{x}| < \delta$.

Another expression that is useful is

$$f(\mathbf{x}) = o(|\mathbf{x}|) \quad \text{which is shorthand for} \quad \lim_{|\mathbf{x}|\to 0} \frac{f(\mathbf{x})}{|\mathbf{x}|} = 0.$$

Informally this means that $f(\mathbf{x})$ vanishes faster than $|\mathbf{x}|$ as $|\mathbf{x}| \to 0$. More generally $f = o(g)$ if $\lim_{|\mathbf{x}|\to 0} |f(\mathbf{x})/g(\mathbf{x})| = 0$, meaning that $f(\mathbf{x})$ vanishes faster than $g(\mathbf{x})$ as $|\mathbf{x}| \to 0$.

label: ex:vp1-ord04

Exercise 1.5

(a) If $f_1 = x$ and $f_2 = y$ show that $f_1 = O(f)$ and $f_2 = O(f)$ where $f(x, y) = (x^2 + y^2)^{\frac{1}{2}}$.

(b) Show that the polynomial $\phi(x, y) = ax^2 + bxy + cy^2$ vanishes to at least the same order as the polynomial $f(x, y) = x^2 + y^2$ at $(0, 0)$. What conditions are needed for ϕ to vanish faster than f as $\sqrt{x^2 + y^2} \to 0$?

1.3 Functions of a real variable

label: sec:vp1-one

1.3.1 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.

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[9], Rudin[10] and by Kolmogorov and Fomin[11].

1.3.2 Continuity and Limits

label: sec:vp1-cont

A *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 1.1 is continuous for $x_1 < x < x_2$. The function $f_2(x)$, depicted by the dashed line, is discontinuous at $x = c$.

label: f:vp1-c01

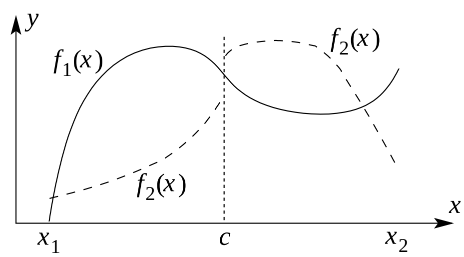

Figure 1.1 Figure showing examples of a continuous function, $f_1(x)$, and a discontinuous function $f_2(x)$.

A function $f(x)$ is continuous at a point $x = a$ if $f(a)$ exists and if, given any arbitrarily small number, we can find a neighbourhood of $x = a$ such that in it $|f(x) - f(a)|$ is always smaller. 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

label: def:vp1-cont1

Definition 1.4
Continuity: a function, f, is continuous at $x = a$ if $f(a)$ is defined and

$$\lim_{x\to a} f(x) = f(a).$$

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

$$\lim_{x\to a_-} f(x) \quad \text{and} \quad \lim_{x\to a_+} f(x),$$

[9] *A Course of Modern Analysis* by E T Whittaker and G N Watson, Cambridge University Press.
[10] *Principles of Mathematical Analysis* by W Rudin (McGraw-Hill)
[11] *Introductory Real Analysis* by A N Kolmogorov and S V Fomin (Dover)

exist *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 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 *removable* discontinuity at $x = a$.

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 course, 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 counter example 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 a commonly occurring piecewise continuous functions that are discontinuous. They are defined by label: eq:vp1-c01

$$H(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0, \end{cases} \qquad \text{and} \quad \text{sgn}(x) = \begin{cases} 1, & x > 0, \\ -1, & x < 0, \end{cases} \qquad \text{sgn}(x) = -1 + 2H(x). \tag{1.4}$$

These functions are discontinous 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 or 1/2.

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

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

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

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

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, but are less easy to recognise.

label: ex:vp1-lim01

Exercise 1.6
A function that is finite and continuous for all x is defined by

$$f(x) = \begin{cases} \dfrac{A}{x^2} + x + B, & 0 \le x \le a, \quad a > 0 \\ \dfrac{C}{x^2} + Dx, & a \le x, \end{cases}$$

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.

label: ex:vp1-lim02

Exercise 1.7
Find the limits of the following functions as $x \to 0$ and $w \to \infty$.
(a) $\dfrac{\sin ax}{x}$, (b) $\dfrac{\tan ax}{x}$, (c) $\dfrac{\sin ax}{\sin bx}$, (d) $\dfrac{3x+4}{4x+2}$, (e) $\left(1 + \dfrac{z}{w}\right)^w$.

For functions of two or more variables, the definition of continuity is essentially the same as for a function of one variable. A function $f(\mathbf{x})$ is continuous at $\mathbf{x} = \mathbf{a}$ if $f(\mathbf{a})$ is defined and

label: eq:vp1-c02

$$\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \lim_{x_1\to a_1} \lim_{x_2\to a_2} \cdots \lim_{x_n\to a_n} f(\mathbf{x}) = f(\mathbf{a}), \tag{1.5}$$

where the order of the limits must be immaterial.

label: ex:vp1-lim03

Exercise 1.8
Determine whether or not the following functions are continuous at the origin.
(a) $f = \dfrac{2xy}{x^2+y^2}$, (b) $f = \dfrac{x^2+y^2}{x^2-y^2}$, (c) $f = \dfrac{2x^2y}{x^2+y^2}$.
Hint: use polar coordinates $x = r\cos\theta$, $y = r\sin\theta$ and consider the limit $r \to 0$.

1.3.3 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 increasing and monotonic decreasing, respectively, on the whole line: the function $f(x) = \sin x$ is monotonic increasing for $-\pi/2 < x < \pi/2$. More precisely, we have,

Definition 1.5
Monotonic functions: A function $f(x)$ is *monotonic* increasing for $a < x < b$ if

$$f(x_1) \le f(x_2) \quad \text{for} \quad a < x_1 < x_2 < b.$$

A monotonic decreasing function is defined in a similar way.

If $f(x_1) < f(x_2)$ for $a < x_1 < x_2 < b$ then $f(x)$ is said to be *strictly monotonic* (increasing) or *strictly increasing*; strictly decreasing functions are defined in the obvious manner.

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 $a \le x \le b$ and $y = f(x)$ the inverse function, $x = f^{-1}(y)$ is continuous for $f(a) \le y \le f(b)$ 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 $x = \pm\sqrt{y}$, $y \ge 0$. These two inverses are often refered 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. label: ex:vp1-mon01

Exercise 1.9

(a) Show that $y = 3a^2x - x^3$ 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 $\cosh^3\phi = (\cosh 3\phi + 3\cosh\phi)/4$.

1.3.4 The derivative

label: sec:vp1-der01

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 1.2: 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

label: f:vp1-der01

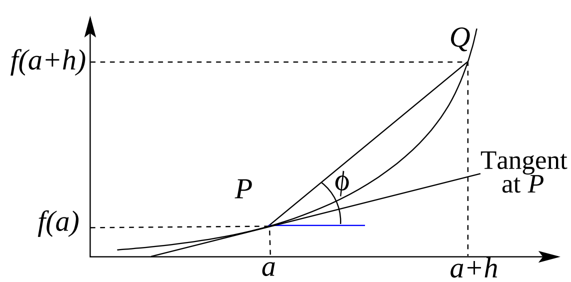

Figure 1.2 Illustration showing the chord PQ and the tangent line at P.

The gradient of the chord PQ is $\tan\phi$ where ϕ is the angle between PQ and the x-axis, and is given by the formula

$$\tan\phi = \frac{f(a+h)-f(a)}{h}.$$

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 $\dfrac{df}{dx}$. Thus we have the formal definition:

label: def:vp1-der01

Definition 1.6
The derivative: A function $f(x)$, defined on an open interval U of the real line, is *differentiable* for $x\in U$ and has the *derivative* $f'(x)$ if

label: eq:vp1-der01

$$f'(x) = \frac{df}{dx} = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}, \tag{1.6}$$

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 1.3 of the order notation shows that a differentiable function satisfies

label: eq:vp1-der02

$$f(x+h) = f(x) + hf'(x) + o(h). \tag{1.7}$$

The formal definition, equation 1.6, of the derivative can be used to derive all its useful properties, but the physical interpretation, illustrated in figure 1.2, 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 1.6, 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 1.6, 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 1.6 the following useful results follow. If $f(x)$ and $g(x)$ are differentable on the same open interval and α and β are constants then

(a) $\displaystyle \frac{d}{dx}\Big(\alpha f(x)+\beta g(x)\Big)=\alpha f'(x)+\beta g'(x),$

(b) $\displaystyle \frac{d}{dx}\Big(f(x)g(x)\Big)=f'(x)g(x)+f(x)g'(x),$ (The product rule)

(c) $\displaystyle \frac{d}{dx}\left(\frac{f(x)}{g(x)}\right)=\frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2},\quad g(x)\neq 0.$ (The quotient rule)

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 1.6, exercise 1.14, 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 1.18, to be given by

label: eq:vp1-der03

$$\frac{dF}{dx}=\frac{df}{dg}\times\frac{dg}{dx}\quad\text{or}\quad F'(x)=f'(g)g'(x). \tag{1.8}$$

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 $f'(g)$ or $f'(g(x))$ does not mean the derivative of $F(x)$; it means the derivative $f'(x)$ with x replaced by g or $g(x)$.

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

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

The derivatives of simple functions, polynomials and trigometric functions for instance, can be deduced from first principles using the definition 1.6: the three rules, given above, and the chain rule can then be used to find the derivative of any function described with finite combinations of these simple functions. A few exercises will make this process clear.

label: ex:vp1-der01

Exercise 1.10
Find the derivative of the following functions

(a) $\sqrt{(a-x)(b+x)}$, (b) $\sqrt{a\sin^2 x+b\cos^2 x}$, (c) $\cos(x^3)\cos x$, (d) x^x.

label: ex:vp1-der02

Exercise 1.11
If $y=\sin x$ for $\pi/2\le x\le\pi/2$ show that $\displaystyle\frac{dx}{dy}=\frac{1}{\sqrt{1-y^2}}$.

label: ex:vp1-der03

Exercise 1.12

(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 $\dfrac{d^2x}{dy^2}$ in terms of $\dfrac{dy}{dx}$ and $\dfrac{d^2y}{dx^2}$.

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

$$f, \quad \frac{df}{dx}, \quad \frac{d^2f}{dx^2}, \quad \frac{d^3f}{dx^3}, \cdots, \frac{d^{n-1}f}{dx^{n-1}}, \quad \frac{d^nf}{dx^n} \cdots,$$

where each member of the sequence is the derivative of the preceeding member,

$$\frac{d^pf}{dx^p} = \frac{d}{dx}\left(\frac{d^{p-1}f}{dx^{p-1}}\right), \quad p = 2, 3, \cdots.$$

The prime notation becomes rather clumsy after the second or third derivative, so the most common alternative is

$$\frac{d^pf}{dx^p} = f^{(p)}(x), \quad p \geq 2,$$

with the conventions $f^{(1)}(x) = f'(x)$ and $f^{(0)}(x) = f(x)$. Care is needed to distinguish between the pth derivative, $f^{(p)}(x)$, and the pth 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 nth derivative is continuous are said to be n-differentiable and to belong to class C^n: the notation $C^n(U)$ means the 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^∞, that is functions, such as $\sin x$, having all derivatives; we shall use the term *sufficiently smooth* for functions that are sufficiently differentiable for all subsequent analysis to work, when more detail is deemed unimportant.

label: ex:vp1-der03b

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

Exercise 1.13

If $f(x)$ is an even (odd) function, show that $f'(x)$ is an odd (even) function.

label: ex:vp1-der04a

Exercise 1.14

Show, from first principles using the limit 1.6, that $\dfrac{d}{dx}\left(\dfrac{1}{f(x)}\right) = -\dfrac{f'(x)}{f(x)^2}$, and that the product rule is true.

label: ex:vp1-der05

Exercise 1.15
Leibniz's rule
If $h(x) = f(x)g(x)$ show that

$$\begin{aligned} h''(x) &= f''(x)g(x) + 2f'(x)g'(x) + f(x)g''(x), \\ h^{(3)}(x) &= f^{(3)}(x)g(x) + 3f''(x)g'(x) + 3f'(x)g''(x) + f(x)g^{(3)}(x), \end{aligned}$$

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 $\displaystyle\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$.

label: ex:vp1-der06

Exercise 1.16
Show that $\displaystyle\frac{d}{dx}\ln(f(x)) = \frac{f'(x)}{f(x)}$ and hence that if

$$p(x) = f_1(x)f_2(x)\cdots f_n(x) \quad \text{then} \quad \frac{p'}{p} = \frac{f_1'}{f_1} + \frac{f_2'}{f_2} + \cdots + \frac{f_n'}{f_n},$$

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.

label: ex:vp1-der07

Exercise 1.17
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 to third-order determinants.

1.3.5 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[12] starting at the point $A = (a, f(a))$ and ending at $B = (b, f(b))$, as shown in figure 1.3.

label: f:vp1-mean01

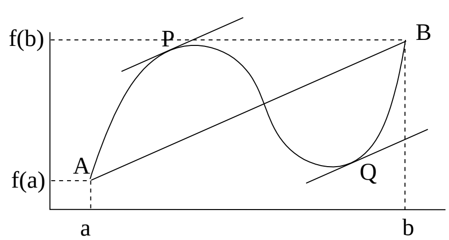

Figure 1.3 Diagram illustrating Cauchy's form of the mean value theorem.

[12] A smooth curve is one along which its tangent changes direction continuously, without abrupt changes.

From this figure it seems plausible that the tangent to the curve must be parallel to the chord AB at least once. That is

label: eq:vp1-mean01

$$f'(x) = \frac{f(b) - f(a)}{b - a} \quad \text{for some } x \text{ in the interval } a < x < b. \tag{1.9}$$

label: eq:vp1-mean01a

Alternatively this may be written in the form

$$f(b) = f(a) + hf'(a + \theta h), \quad h = b - a. \tag{1.10}$$

where θ is a number in the interval $0 < \theta < 1$, which is normally unknown. Note that this shows that between zeros of a continuous function there is at least one point at which the derivative is zero. Equation 1.9 can be proved and is enshrined in the following theorem

label: ther:vp1-mean01

Theorem 1.1
The Mean Value Theorem (Cauchy's form). If $f(x)$ and $g(x)$ are real and differentiable for $a \le x \le b$, then there is a point u inside the interval at which

label: eq:vp1-mean02

$$\Big(f(b) - f(a)\Big)\, g'(u) = \Big(g(b) - g(a)\Big)\, f'(u), \quad a < u < b. \tag{1.11}$$

By putting $g(x) = x$, equation 1.9 follows.

A similar idea may be applied to integrals. In figure 1.4 is shown a typical continuous function, $f(x)$, which attains its smallest and largest values, S and L respectively, on the interval $a \le x \le b$.

label: f:vp1-mean02

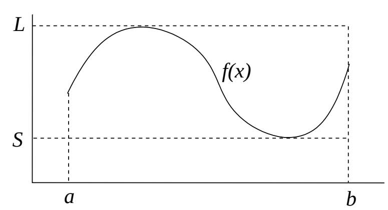

Figure 1.4 Diagram showing the upper and lower bounds of $f(x)$ used to bound the integral.

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

$$(b - a)S \le \int_a^b dx\, f(x) \le (b - a)L.$$

label: eq:vp1-mean03

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

$$\int_a^b dx\, f(x) = (b - a)f(\xi) \quad \text{for some } \xi \in (a, b). \tag{1.12}$$

label: ther:vp1-mean02

This observation is made rigorous in the following theorem.

Theorem 1.2
The Mean Value theorem (integral form). If, on the closed interval $a \le x \le b$, $f(x)$ is continuous and $\phi(x) \ge 0$ then there is an ξ satisfying $a \le \xi \le b$ such that

label: eq:vp1-mean04

$$\int_a^b dx\, f(x)\phi(x) = f(\xi) \int_a^b dx\, \phi(x). \tag{1.13}$$

If $\phi(x) = 1$ relation 1.12 is regained.

label: ex:vp1-mean01

Exercise 1.18
The chain rule
In this exercise the Mean Value Theorem is used to derive the chain rule, equation 1.8, for the derivative of $F(x) = f(g(x))$.

Use the mean value theorem to show that

$$F(x+\epsilon) - F(x) = f\left(g(x) + \epsilon g'(x+\epsilon\theta)\right) - f(g(x))$$

and that

$$f\left(g(x) + \epsilon g'(x+\epsilon\theta)\right) = f(g(x)) + \epsilon g'(x+\epsilon\theta)\, f'(g+\epsilon\phi g')$$

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

$$\frac{F(x+\epsilon) - F(x)}{\epsilon} = f'(g+\epsilon\phi g')\, g'(x+\epsilon\theta),$$

and by taking the limit $\epsilon \to 0$ derive equation 1.8.

label: ex:vp1-mean02

Exercise 1.19
Use the integral form of the mean value theorem to evaluate the limits,

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

1.3.6 Partial Derivatives

label: sec:vp1-part

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 1.6 (page 16).

label: def:vp1-der02

Definition 1.7
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 derivative with respect to x and y are, repectively

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

label: ex:vp1-part01

Exercise 1.20

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

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

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

label: eq:vp1-part01

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx}(x, y) \quad \text{and} \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy}(x, y). \tag{1.14}$$

label: eq:vp1-part02

But now we also have the mixed derivatives

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \quad \text{and} \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right). \tag{1.15}$$

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

label: eq:vp1-part03

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}. \tag{1.16}$$

In terms of 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, see equation 1.5 (page 14).

Similarly, differentiating p times with respect to x and q times with respect to y, in any order, gives the same nth order derivative,

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

provided all the nth derivatives are continuous.

label: ex:vp1-part02

Exercise 1.21

If $\Phi(x, y) = \exp(-x^2/y)$ show that Φ satisfies the equations

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

label: ex:vp1-part03

Exercise 1.22
Show that $u = x^2 \sin(\ln y)$ satisfies the equation $2y^2 \dfrac{\partial^2 u}{\partial y^2} + 2y \dfrac{\partial u}{\partial y} + x \dfrac{\partial u}{\partial x} = 0$.

The generalisation of these ideas to functions of the n variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is straightforward: the partial derivative of $f(\mathbf{x})$ with respect to x_k is defined to be label: eq:vp1-part04

$$\frac{\partial f}{\partial x_k} = \lim_{h\to 0} \frac{f(x_1, x_2, \cdots, x_{k-1}, x_k + h, x_{k+1}, \cdots, x_n) - f(x_1, x_2, \ldots, x_n)}{h}. \qquad (1.17)$$

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.

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 label: eq:vp1-part05

$$F(t) = f(x_1(t), x_2(t), \cdots, x_n(t)). \qquad (1.18)$$

The derivative of $F(t)$ is given by the relation label: eq:vp1-part06

$$\frac{dF}{dt} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \frac{dx_k}{dt}. \qquad (1.19)$$

Normally, we write $f(t)$ rather than use a different symbol $F(t)$, and the left hand side of the above equation is written $\dfrac{df}{dt}$. 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

$$\begin{aligned} F(t+\epsilon) &= f(x(t+\epsilon), y(t+\epsilon)) \\ &= f\Big(x(t) + \epsilon x'(t+\theta\epsilon), y(t) + \epsilon y'(t+\phi\epsilon)\Big), \quad 0 < \theta, \phi < 1, \end{aligned}$$

where we have used the mean value theorem, equation 1.10. Write the right hand side in the form

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

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

Equation 1.19 is used in chapter 3 to derive one of the most important results in the course: if the dependence of $\mathbf{x}$ upon t is linear and $F(t)$ has the form

$$F(t) = f(\mathbf{x} + t\mathbf{h}) = f(x_1 + th_1, x_2 + th_2, \cdots, x_n + th_n)$$

where the vector $\mathbf{h}$ is constant and the variable x_k has been replaced by $x_k + th_k$, for

label: eq:vp1-part06b

all k. Since $\frac{d}{dt}(x_k + th_k) = h_k$, equation 1.19 becomes

$$\frac{dF}{dt} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} h_k. \tag{1.20}$$

This result will also be used in section 1.3.9 to derive the Taylor series for several variables.

A variant of equation 1.18, which frequently occurs in the Calculus of Variations, is the case where $f(\mathbf{x})$ depends explicity upon the variable t, so this equation becomes

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

label: eq:vp1-part06a

and then equation 1.19 acquires an additional term,

$$\frac{dF}{dt} = \frac{\partial f}{\partial t} + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \frac{dx_k}{dt}. \tag{1.21}$$

The right hand side of equations 1.19 and 1.21 depend upon both $\mathbf{x}$ and t, but because $\mathbf{x}$ 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 $\mathbf{x}$ and t, separately, is important: for instance we often require expressions like

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

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

$$\begin{aligned}\frac{\partial}{\partial x_1}\left(\frac{dF}{dt}\right) &= \frac{\partial^2 f}{\partial x_1 \partial t} + + \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_1 \partial x_k}\frac{dx_k}{dt} \\ &= \frac{d}{dt}\left(\frac{\partial F}{\partial x_1}\right),\end{aligned}$$

label: ex:vp1-part05

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

Exercise 1.23
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

$$\begin{aligned} \frac{\partial}{\partial y}\left(\frac{df}{dt}\right) &= \frac{dx}{dt} - 2y - 2t\frac{dy}{dt} = 2t\left(1 - 4t^2\right), \\ \frac{d}{dt}\left(\frac{\partial f}{\partial y}\right) &= \frac{d}{dt}(x - 2ty) = \frac{dx}{dt} - 2y - 2t\frac{dy}{dt}. \end{aligned}$$

label:
ex:vp1-part05a

Exercise 1.24
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

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

label:
ex:vp1-part06

Exercise 1.25
Euler's formula for homogeneous functions

(a) If $f(x,y)$ is a function of two variables with the property $f(\lambda x, \lambda y) = \lambda^p f(x,y)$, for any constant λ and any real number p, show that

$$pf(x,y) = xf_x(x,y) + yf_y(x,y).$$

Hint: use the total derivative formula 1.19 and differentiate with respect to λ.

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

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

1.3.7 Implicit functions

label:
sec:vp1-imp01

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 1.5.

label:
f:vp1-imp01

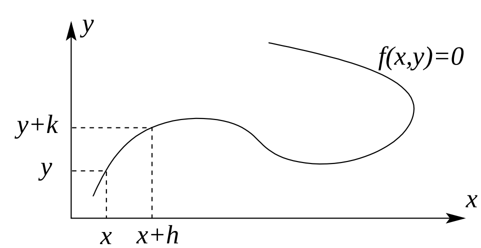

Figure 1.5 Diagram showing a typical curve defined by an equation of the form $f(x,y) = 0$.

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 1.5. 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 1.29, 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 1.5, so $F(x)=f(x,y(x))$ is a function of x only and we may use the chain rule 1.21 to differentiate with respect to x. This gives

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

label: eq:vp1-imp02a

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

$$\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{dy}{dx}=0 \quad \text{or} \quad \frac{dy}{dx}=-\frac{f_x}{f_y}. \tag{1.22}$$

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

label: eq:vp1-imp02b

$$\frac{\partial f}{\partial x}\frac{dx}{dy}+\frac{\partial f}{\partial y}=0 \quad \text{or} \quad \frac{dx}{dy}=-\frac{f_y}{f_x}. \tag{1.23}$$

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 $f(x,y(x))=0$. We shall give a restricted version of it here.

Theorem 1.3
Implicit Function Theorem: Suppose that $f:U\to R$ is a function with continuous partial derivatives defined in an open set $U\subseteq R^2$. 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 $f(x,y(x))=0$. The function $f:I\to J$ is continuous, differentiable, with the derivative given by equation 1.22.

label: ex:vp1-imp03

Exercise 1.26
In the case $f(x,y)=y-g(x)$ show that equations 1.22 and 1.23 leads to the relation

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

label: ex:vp1-imp01

Exercise 1.27

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

label: ex:vp1-imp02

Exercise 1.28

If $x - y + \sin(xy) = 0$ determine the values of $y'(0)$ and $y''(0)$.

label: ex:vp1-imp04

Exercise 1.29

Show that the differential equation

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

has a solution defined by the equation

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

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

label: eq:vp1-imp03

$$f_k(\mathbf{x}, \mathbf{t}) = 0, \quad k = 1,\, 2, \cdots, n \tag{1.24}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{t} = (t_1, t_2, \ldots, t_m)$. These n equations have a unique solution for each x_k in terms of $\mathbf{t}$, $x_k = g_k(\mathbf{t})$, $k = 1,\, 2, \cdots, n$, in the neighbourhood of $(\mathbf{x}_0, \mathbf{t}_0)$ provided that at this point the derivatives $\partial f/\partial x_k$, exist and that the determinant

label: eq:vp1-imp04

$$J = \begin{vmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{vmatrix} \tag{1.25}$$

is not zero. Furthermore all the functions $g_k(\mathbf{t})$ have continuous first derivatives. The determinant J is named the *Jacobian determinant*.

label: ex:vp1-imp05

Exercise 1.30

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.

1.3.8 Taylor series for one variable

label: sec:vp1-tay

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 1.4
Taylor's Theorem: If $f(x)$ is a function defined on $x_1 \le x \le x_2$ such that $f^{(n)}(x)$ is continuous for $x_1 \le x \le x_2$ and $f^{(n+1)}(x)$ exists for $x_1 < x < x_2$, then if $a \in [x_1, x_2]$ for every $x \in [x_1, x_2]$

label: eq:vp1-tay01

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a) + R_{n+1}. \quad (1.26)$$

label: eq:vp1-tay02

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

$$R_{n+1} = \frac{(x-a)^{n+1}}{(n+1)!}f^{(n+1)}(a+\theta h) \quad \text{for some } 0 < \theta < 1 \text{ and } h = x-a. \quad (1.27)$$

If all derivatives of $f(x)$ are continuous for $x_1 \le x \le x_2$, 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

label: eq:vp1-tay03

$$f(x) = \sum_{k=0}^{\infty} \frac{(x-a)^k}{k!} f^{(k)}(a). \quad (1.28)$$

The infinite series 1.28 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 1.26 as $n \to \infty$, because there are cases then the infinite series on the right hand side of equation 1.28 *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 ξ, 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 $f(x)$ 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 course and we shall assume that the Taylor series derived from a function converges to it in some interval.

The series 1.28 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 1.28, 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

label: eq:vp1-tay03a

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a). \quad (1.29)$$

The series 1.28 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 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 repectively, then the Taylor series of $\alpha f(x) + \beta g(x)$, $f(x)g(x)$ $f(g(x))$ and $g(f(x))$ exist and have the radius of convergence $\min(r_f, r_g)$: 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 nth 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 $\tanh x$,

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + O(z^5) \quad \text{and} \quad \tanh x = x - \frac{x^3}{3} + \frac{2x^5}{15} + O(x^7),$$

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

$$\begin{aligned} \ln(1+\tanh x) &= \left[x - \frac{x^3}{3} + O(x^5)\right] - \left[\frac{x^2}{2}\left(1 - \frac{x^2}{3} + \cdots\right)^2\right] + \frac{x^3}{3} - \frac{x^4}{4} + O(x^5) \\ &= x - \frac{x^2}{2} + \frac{x^4}{12} + O(x^5). \end{aligned}$$

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

For $|x-a| > r_c$ the infinite sum 1.28 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 label: eq:vp1-tay04

$$r_c = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| \quad \text{or} \quad r_c = \lim_{n\to\infty}|a_n|^{-1/n} \quad \text{where} \quad a_k = \frac{f^{(k)}(a)}{k!}. \tag{1.30}$$

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

$$\left|\frac{n!}{f^{(n)}(a)}\right|^{1/n} = r_c\Big(1 + O(1/n)\Big) \quad \text{so that} \quad \frac{n!}{f^{(n)}(a)} = A r_c^n \Big(1 + O(1/n)\Big)$$

for some constant A. Then the nth term of the series behaves as $((a-x)/r_c)^n$, and decreases rapidly with increasing n provided $|a-x| < r_c$ 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 $|a-x|$ 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 later in the course.

The first practical problem is that the remainder term, equation 1.27, depends upon θ, the value of which is unknown. Hence R_n cannot be computed; also, it is normally difficult to estimate.

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

label: eq:vp1-tay05

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \Big(1+O(1/n)\Big) \tag{1.31}$$

label: eq:vp1-tay06

we can approximate the first part of this product by,

$$\left|\frac{(x-a)^n}{n!}\right| \simeq \frac{1}{\sqrt{2\pi n}}\left(\frac{e|x-a|)}{n}\right)^n = g_n. \tag{1.32}$$

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 1.6 we plot a graph of the values of $\log(g_n)$, that is the logarithm to the base 10, for $x-a=10$.

label: f:vp1-tay01

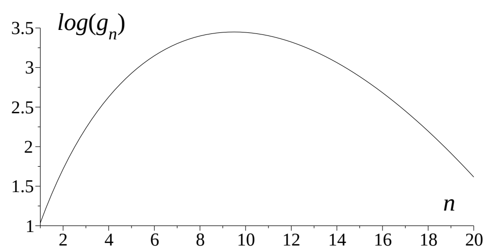

Figure 1.6 Graph showing the value of $\log(g_n)$, equation 1.32, for $x-a=10$. For clarity we have joined the points with a continuous line.

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 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 examples.

label: eq:vp1-tay07

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

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + (-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!} + \cdots, \tag{1.33}$$

which is derived in exercise 1.31.

Note that only odd powers occur and because $\sin x$ is an odd function and also that the radius of convergence is infinite. In figure 1.7 we show graphs of this series, truncated at x^{2n-1} with $n = 1$, 4, 8 and 15 for $0 < x < 4\pi$.

label: f:vp1-tay02

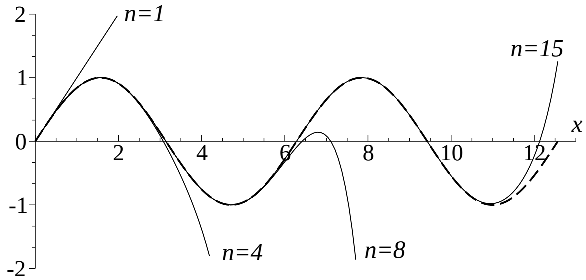

Figure 1.7 Graph comparing the Taylor polynomials, of order n, for the sine function with the exact function, the dashed line.

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 1.6. 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 of this behaviour using a computer having finite accuracy there is value of x beyond which the Taylor series for $\sin x$ gives incorrect values, despite the fact that formally it converges for all x.

label: ex:vp1-tay01

Exercise 1.31
Exponentional 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

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{(-1)^n x^{2n}}{(2n)!} + \cdots , \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{(-1)^n x^{2n-1}}{(2n-1)!} + \cdots . \end{aligned}$$

label: ex:vp1-tay02

Exercise 1.32
Binomial expansion
Show that the Taylor series of $(1 + a)^a$ 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.

label: ex:vp1-tay03

Exercise 1.33

If $f(x) = \tan x$ find the first three derivatives to show that $\tan x = x + \frac{1}{3}x^3 + O(x^5)$.

label: ex:vp1-tay04

Exercise 1.34

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,

$$\ln(1+x) = \int_0^x dt\, \frac{1}{1+t},$$

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

$$\ln\left(\frac{1+x}{1-x}\right) = 2\left(x + \frac{x^3}{3} + \cdots + \frac{x^{2n-1}}{2n-1} + \cdots\right).$$

label: ex:vp1-tay05

Exercise 1.35

Use the definition $\tan^{-1} x = \int_0^x dt\,(1+t^2)^{-1}$ to show that for $|x| < 1$,

$$\tan^{-1} x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1}.$$

label: ex:vp1-tay06

Exercise 1.36

Show that

$$\ln(1 + \sinh x) = x - \frac{x^2}{2} + \frac{x^3}{2} - \frac{5x^4}{12} + O(x^5).$$

label: ex:vp1-tay07

Exercise 1.37

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 1.15 (page 19) to differentiate the equation n times.

1.3.9 Taylor series for several variables

label: sec:vp1-tays

The Taylor expansion of a function $f : R^n \to R$ is trivially derived from the Taylor expansion of a function of one variable, using the the chain rule, in particular the version described in equation 1.21 (page 24). The only difficulty is that the algebra very quickly becomes unwieldy with increasing order.

We require the expansion of $f(\mathbf{x})$ about $\mathbf{x} = \mathbf{a}$, so we need to represent $f(\mathbf{a}+\mathbf{h})$ as some sort of power series in $\mathbf{h}$. To this end, define a function of the single variable t by the relation

$$F(t) = f(\mathbf{a}+t\mathbf{h}) \quad \text{so} \quad F(0) = f(\mathbf{a}),$$

and $F(t)$ gives values of $f(\mathbf{x})$ on the straight line joining $\mathbf{a}$ to $\mathbf{a}+\mathbf{h}$. The Taylor series of $F(t)$ about $t = 0$ is, on using equation 1.26 (page 28),

label: eq:vp1-tays01

$$F(t) = F(0) + tF'(0) + \frac{t^2}{2!}F''(0) + \cdots + \frac{t^n}{n!}F^{(n)}(0) + R_{n+1}, \tag{1.34}$$

which we assume to exist for $|t| \le 1$. Now we need only express the derivatives $F^{(n)}(0)$ in terms of the partial derivatives of $f(\mathbf{x})$. Equation 1.21 (page 24) gives

$$F'(0) = \sum_{k=1}^{n} f_{x_k}(\mathbf{a})h_k.$$

Hence the first-order Taylor polynomial is

label: eq:vp1-tays02

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \sum_{k=1}^{n} h_k f_{x_k}(\mathbf{a}) + R_2 = f(\mathbf{a}) + \mathbf{h}\cdot\frac{\partial f}{\partial \mathbf{a}} + R_2, \tag{1.35}$$

where R_2 is the remainder term which is second order in $\mathbf{h}$. Here we have introduced the notation $\partial f/\partial\mathbf{x}$ for the vector function,

$$\frac{\partial f}{\partial \mathbf{x}} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n}\right) \quad \text{with the scalar product} \quad \mathbf{h}\cdot\frac{\partial f}{\partial \mathbf{x}} = \sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k}.$$

The remaining terms are derived in exactly the same manner, but the algebra quickly becomes cumbersome.

For the second derivative we simply differentiate equation 1.21 again,

$$F''(t) = \sum_{k=1}^{n} h_k \frac{d}{dt} f_{x_k}(\mathbf{a}+t\mathbf{h}) = \sum_{k=1}^{n} h_k \left(\sum_{i=1}^{n} h_i f_{x_k x_i}(\mathbf{a}+t\mathbf{h})\right).$$

Hence, we have

label: eq:vp1-tays03

$$\begin{aligned} F''(0) &= \sum_{k=1}^{n} h_k \sum_{i=1}^{n} h_i f_{x_k x_i}(\mathbf{a}), \\ &= \sum_{k=1}^{n} h_k^2 f_{x_k x_k}(\mathbf{a}) + 2\sum_{k=1}^{n-1}\sum_{i=k+1}^{n} h_k h_i f_{x_k x_i}(\mathbf{a}), \end{aligned} \tag{1.36}$$

where the second relation comprises fewer terms because the mixed derivative rule has been used. Similarly, we see that

$$F^{(3)}(0) = \sum_{k=1}^{n} h_k \left(\sum_{i=1}^{n} h_i \left(\sum_{j=1}^{n} h_j f_{x_k\, x_i\, x_j}(\mathbf{a}) \right) \right).$$

label: eq:vp1-tays04

Thus to second order the Taylor series is

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \sum_{k=1}^{n} h_k f_{x_k}(\mathbf{a}) + \frac{1}{2!} \left(\sum_{k=1}^{n} h_k \sum_{i=1}^{n} h_i f_{x_k\, x_i}(\mathbf{a}) \right) + R_3. \tag{1.37}$$

The remainder term for the series for $F(t)$ is given in equation 1.27 and so for the series for f we have

$$R_{n+1} = \frac{1}{(n+1)!} F^{(n+1)}(\theta) \quad \text{for some} \quad 0 < \theta < 1.$$

Because the high order derivatives are so cumbersome and for the practical reasons discussed in section 1.3.8, in particular figure 1.7 (page 31), Taylor series for two or more variables are rarely used beyond the second-order term. This term, however, is important for the classification of stationary points, considered in chapter 6.

label: ex:vp1-tays01

Exercise 1.38
Find the Taylor expansions about $x = y = 0$, up to and including the second-order terms, of the functions
(a) $f(x, y) = \sin x \sin y$, (b) $f(x, y) = \sin\left(x + e^{-y} - 1\right)$.

1.3.10 L'Hospital's rule

label: eq:vp1-hop01

Ratios of functions occur frequently and if

$$R(x) = \frac{f(x)}{g(x)} \tag{1.38}$$

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

label: eq:vp1-hop01a

$$R(x) = \frac{\sin x}{x} \tag{1.39}$$

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 and its extensions.

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 $|\epsilon|$,

$$R(a+\epsilon) = \frac{f(x+\epsilon)}{g(a+\epsilon)} = \frac{\epsilon f'(a) + O(\epsilon^2)}{\epsilon g'(a) + O(\epsilon^2)} = \frac{f'(a)}{g'(a)} + O(\epsilon) \quad \text{provided} \quad g'(a) \neq 0.$$

Hence, on taking the limit $\epsilon \to 0$, we obtain the result given by the following theorem.

Theorem 1.5
L'Hospital's rule. Suppose that $f(x)$ and $g(x)$ are real and differentiable for $-\infty \le a < x < b \le \infty$. If

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0 \quad \text{or} \quad \lim_{x\to a} g(x) = \infty$$

then

label: eq:vp1-hop02

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}. \tag{1.40}$$

More generally if $f^{(k)}(a) = g^{(k)}(a) = 0$, $k = 0, 1, \cdots, n-1$ and $g^{(n)}(a) \ne 0$ then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f^{(n)}(x)}{g^{(n)}(x)}.$$

Consider the function defined by equation 1.39; at $x = 0$ L'Hospital's rule gives

$$R(0) = \lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = 1.$$

label: ex:vp1-hop01

Exercise 1.39
Find the values of the following limits:

$$\text{(a)} \lim_{x\to a} \frac{\cosh x - \cosh a}{\sinh x - \sinh a}, \quad \text{(b)} \lim_{x\to 0} \frac{\sin x - x}{x\cos x - x}, \quad \text{(c)} \lim_{x\to 0} \frac{3^x - 3^{-x}}{2^x - 2^{-x}}.$$

label: ex:vp1-hop02

Exercise 1.40
(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 $\lim_{x\to a} \frac{f(x)}{g(x)} = A$.

1.3.11 Integration

label: sec:vp1-int

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 37, 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 course, you should be familiar with its contents. The important idea, needed in chapter 3, is that of differentiating with respect to a parameter, or 'differentiating under the integral sign' described in equation 1.47 (page 39).

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 $a \le x \le b$, 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 1.8.

label: f:vp1-int01

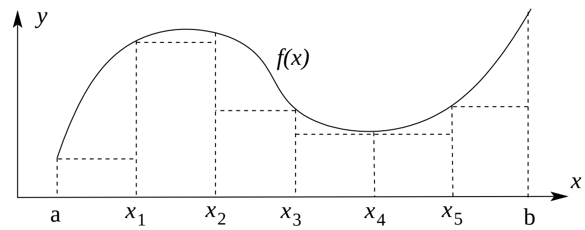

Figure 1.8 Diagram showing how the area under the curve $y = f(x)$ may be approximated by a set of rectangles. The intervals $x_k - x_{k-1}$ need not be the same length.

In general the closed interval $a \le x \le b$ 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 1.8 $n = 6$ and the spacings are equal. On each interval we construct a rectangle: on the kth 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

label: eq:vp1-int01

$$\sum_{k=1}^{n} (x_k - x_{k-1})\, f(l_k) \le A \le \sum_{k=1}^{n} (x_k - x_{k-1})\, f(h_k). \tag{1.41}$$

This type of approximation underlies the simplest numerical methods of approximating integrals and, as will be seen in chapter 3, 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

label: eq:vp1-int02

$$\int_a^b dx\, f(x) \quad \text{or} \quad \int_a^b f(x)\, dx. \tag{1.42}$$

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 continuous 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: $\int_a^b dx\, \Big(\alpha f(x) + \beta g(x)\Big) = \alpha \int_a^b dx\, f(x) + \beta \int_a^b dx\, g(x)$, where α and β are real or complex numbers.

P V: $\left| \int_a^b dx\, f(x) \right| \le \int_a^b dx\, |f(x)|$. This is the analogue of the finite sum inequality $\left| \sum_{k=1}^{n} a_k \right| \le \sum_{k=1}^{n} |a_k|$, where a_k, $k = 1, 2, \cdots, n$, are a set of complex numbers or functions.

P VI: The *Hölder inequality*: if $\frac{1}{p} + \frac{1}{q} = 1$, $p > 1$ and $q > 1$ then

$$\left| \int_a^b dx\, f(x)g(x) \right| \le \left(\int_a^b dx\, |f(x)|^p \right)^{1/p} \left(\int_a^b dx\, |g(x)|^q \right)^{1/q},$$

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. This is the analogue of the finite sum inequality

$$\sum_{k=1}^{n} |a_k b_k| \le \left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p} \left(\sum_{k=1}^{n} |b_k|^q \right)^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

with equality if and only if $|a_n|^p |b_n|^{-q}$ and $\arg(a_n b_n)$ are independent of n (or $a_k = 0$ for all k or $b_k = 0$ for all k).

When $p = q = 2$ this becomes the Cauchy-Schwarz inequality, which is sometimes named the Cauchy inequality and sometimes the Schwarz inequality.

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

label: eq:vp1-int03

$$\int_a^\infty dx\, f(x) = \lim_{b\to\infty} \left(\int_a^b dx\, f(x) \right), \tag{1.43}$$

assuming the limit exists. There are similar definitions for

$$\int_{-\infty}^b dx\, f(x) \quad \text{and} \quad \int_{-\infty}^\infty dx\, f(x),$$

however, it should be noted that the limit

$$\lim_{a\to\infty} \int_{-a}^a dx\, f(x) \quad \text{may exist, but the limit} \quad \lim_{a\to\infty} \lim_{b\to\infty} \int_{-b}^a dx\, f(x)$$

may not. An example is $f(x) = x/(1+x^2)$ for which

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

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 1.43. If $f(x) \neq 0$ for some $X > 0$, the limit exist 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 1.50, (page 41).

If the integrand is oscillatory, however, 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.

Theorem 1.6
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, mentioned above, which ensures convergence can also 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

label: eq:vp1-int04

$$\int_a^b dx\, f(x) = \lim_{\epsilon\to 0_+} \int_{a+\epsilon}^b dx\, f(x), \tag{1.44}$$

provided the limit exists. For functions unbounded at an interior point the natural extension to P III is used. As a general rule, provided $|f(x)|$ tends to infinity slower $|x-a|^\beta$, $\beta > -1$, 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.

The evaluation of integrals of any complexity 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: label: eq:vp1-int05

$$\int_a^b dx\, u\frac{dv}{dx} = \Big[uv\Big]_a^b - \int_a^b dx\, \frac{du}{dx}v. \tag{1.45}$$

The second method is to change variables: label: eq:vp1-int06

$$\int_a^b dx\, f(x) = \int_A^B dt\, \frac{dx}{dt} f(g(t)) = \int_A^B dt\, g'(t) f(g(t)), \tag{1.46}$$

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, so making the equation easier to remember. The geometric significance of this formula is simply that the small element of length δx, 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 label: eq:vp1-int07

$$\frac{d}{du}\int_{a(u)}^{b(u)} dx\, f(x,u) = f(b,u)\frac{db}{du} - f(a,u)\frac{da}{du} + \int_{a(u)}^{b(u)} dx\, \frac{\partial f}{\partial u} \tag{1.47}$$

provided $a(u)$ and $b(u)$ are differentiable and $f_u(x, u)$ is a continuous function of both variables, see equation 1.5. 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

$$I(u) = \int_0^\infty dx\, e^{-xu}, \quad u > 0.$$

The derivatives are

$$I'(u) = -\int_0^\infty dx\, xe^{-xu} \quad \text{and, in general,} \quad I^{(n)}(u) = (-1)^n \int_0^\infty dx\, x^n e^{-xu}.$$

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

$$\int_0^\infty dx\, x^n e^{-xu} = \frac{n!}{u^{n+1}}.$$

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. label: ex:vp1-int01a

Exercise 1.41

(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 $\int_{-a}^{a} dx\, f(x) = 2\int_{0}^{a} dx\, f(x)$.

label: ex:vp1-int01b

Exercise 1.42
Show that, if $\lambda > 0$, the value of the integral $I(\lambda) = \int_0^\infty dx\, \frac{\sin \lambda x}{x}$ is independent of λ. How are the values of $I(\lambda)$ and $I(-\lambda)$ related?

label: ex:vp1-int01

Exercise 1.43
Use integration by parts to evaluate the following indefinite integrals.

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

label: ex:vp1-int02

Exercise 1.44
Evaluate the following integrals

$$\text{(a)} \quad \int_0^{\pi/4} dx\, \sin x\, \ln(\cos x), \quad \text{(b)} \quad \int_0^{\pi/4} dx\, x \tan^2 x, \quad \text{(c)} \quad \int_0^1 dx\, x^2 \sin^{-1} x.$$

label: ex:vp1-int03

Exercise 1.45

If $I_n = \int_0^x dt\, t^n e^{at}$ use integration by parts to show that $aI_n = x^n e^{ax} - nI_{n-1}$ and deduce that

$$I_n = n! e^{ax} \sum_{k=0}^{n} \frac{(-1)^{n-k}}{a^{n-k+1} k!} x^k - \frac{(-1)^n n!}{a^{n+1}}$$

label: ex:vp1-int03a

Exercise 1.46
(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 $I = \pi/4$.

label: ex:vp1-int04

Exercise 1.47
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,

$$\int_0^\pi \frac{dx}{(a + b\cos x)^2}, \int_0^\pi \frac{dx}{(a + b\cos x)^3}, \int_0^\pi dx\, \frac{\cos x}{(a + b\cos x)^2}, \int_0^\pi dx\, \ln(a + b\cos x).$$

label:
ex:vp1-int05

Exercise 1.48
Prove that $y(t) = \frac{1}{\omega}\int_a^t dx\, f(x) \sin\omega(t-x)$ is the solution of the differential equation

$$\frac{d^2y}{dt^2} + \omega^2 y = f(t), \quad y(a) = 0, \quad y'(a) = 0.$$

label:
ex:vp1-int06

Exercise 1.49

(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 1.6 (page 16), 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 $F'(u)$.

label:
ex:vp1-int07

Exercise 1.50
Find the limits as $X \to \infty$ of the following integrals

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

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

1.4 Miscellaneous exercises

Limits

label: ex:vp1-01e

Exercise 1.51

Find, using first principles, the following limits

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

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

where a is a real number.

Inverse functions

label: ex:vp1-71e

Exercise 1.52

Show that the inverse functions of $y=\cosh x$, $y=\sinh x$ and $y=\tanh x$, for $x>0$ are, respectively

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

label: ex:vp1-72e

Exercise 1.53

The function $y=\sin x$ may be defined to be the solution of the differential equation

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

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

$$\frac{d^2x}{dy^2}=y\left(\frac{dx}{dy}\right)^3 \quad \text{which gives} \quad x(y)=\sin^{-1}y=\int_0^y du\frac{1}{\sqrt{1-u^2}}.$$

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 $z=dx/dy$.

Derivatives

label: ex:vp1-03e

Exercise 1.54

Find the derivatives of $y(x)$ where

$$\text{(a)}\ y=f(x)^{g(x)}, \quad \text{(b)}\ y=\sqrt{\frac{p+x}{p-x}}\sqrt{\frac{q+x}{q-x}}, \quad \text{(c)}\ y^n=x+\sqrt{1+x^2}.$$

label: ex:vp1-04e

Exercise 1.55

If $y=\sin(a\sin^{-1}x)$ show that $(1-x^2)y''-xy'+a^2y=0$.

label: ex:vp1-05e

Exercise 1.56

If $y(x)$ satisfies the equation $(1-x^2)\dfrac{d^2y}{dx^2} - 2x\dfrac{dy}{dx} + \lambda y = 0$, where λ is a constant, show that changing the independent variable, x, to θ where $x = \cos\theta$ changes this to

$$\frac{d^2y}{d\theta^2} + \cot\theta\frac{dy}{d\theta} + \lambda y = 0.$$

label:
ex:vp1-06e

Exercise 1.57

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.

Partial derivatives

label:
ex:vp1-10e

Exercise 1.58

If $z = f(x+ay) + g(x-ay) - \dfrac{x}{2a^2}\cos(x+ay)$ where $f(u)$ and $g(u)$ are arbitrary functions of a single variable and a is a constant, prove that

$$a^2\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} = \sin(x+ay).$$

label:
ex:vp1-11e

Exercise 1.59

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).

label:
ex:vp1-12e

Exercise 1.60

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 $\dfrac{1}{x}\dfrac{\partial u}{\partial x} + \dfrac{1}{y}\dfrac{\partial u}{\partial y} + \dfrac{1}{z}\dfrac{\partial u}{\partial z} = \dfrac{1}{u}$.

Implicit functions

label:
ex:vp1-20e

Exercise 1.61

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.

label: ex:vp1-21e

Exercise 1.62

Prove that the equation $x \cos xy = 0$ has a unique solution, $y(x)$, near the point $(1, \frac{\pi}{2})$, and find its first and second derivatives.

label: ex:vp1-22e

Exercise 1.63

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 $(a2^{1/3}, a4^{1/3})$.

Taylor series

label: ex:vp1-31e

Exercise 1.64

By sketching the graphs of $y = \tan x$ and $y = 1/x$ 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 $x_n = n\pi + \dfrac{1}{n\pi}$.

label: ex:vp1-32e

Exercise 1.65

Determine the constants a and b such that $(1 + a \cos 2x + b \cos 4x)/x^4$ is finite at the origin.

label: ex:vp1-33e

Exercise 1.66

Find the Taylor series, to 4th order, of the following functions:

(a) $\ln \cosh x$, (b) $\ln(1 + \sin x)$, (c) $e^{\sin x}$, (d) $\sin^2 x$.

Mean value theorem

label: ex:vp1-41e

Exercise 1.67

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

label: ex:vp1-42e

Exercise 1.68

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

label: ex:vp1-51e

Exercise 1.69
Show that $\lim_{x\to 1} \frac{\sin\ln x}{x^5 - 7x^3 + 6} = -\frac{1}{16}$.

label: ex:vp1-52e

Exercise 1.70
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

label: ex:vp1-61e

Exercise 1.71
Using differentiation under the integral sign show that

$$\int_0^\infty dx\, \frac{\tan^{-1}(ax)}{x(1+x^2)} = \frac{1}{2}\pi\ln(1+a).$$

label: ex:vp1-62e

Exercise 1.72
Prove that, if $|a| < 1$

$$\int_0^{\pi/2} dx\, \frac{\ln(1+\cos\pi a\cos x)}{\cos x} = \frac{\pi^2}{8}(1-4a^2).$$

label: ex:vp1-63e

Exercise 1.73
If $f(x) = (\sin x)/x$, show that $\int_0^{\pi/2} dx\, f(x)f(\pi/2 - x) = \frac{2}{\pi}\int_0^\pi dx\, f(x)$.

Hint: the identity $\sin 2x = 2\sin x\cos x$ is useful.

label: ex:vp1-64e

Exercise 1.74
Use the integral definition

$$\tan^{-1} x = \int_0^x dt\, \frac{1}{1+t^2} \quad \text{to show that for } x > 0 \quad \tan^{-1}(1/x) = \int_x^\infty dt\, \frac{1}{1+t^2}$$

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

label: ex:vp1-65e

Exercise 1.75
Determine the values if x that make $g'(x) = 0$ if $g(x) = \int_x^{2x} dt\, f(t)$ if

(a) $f(t) = e^t$, and (b) $f(t) = (\sin t)/t$.

label: ex:vp1-66e

Exercise 1.76
If $f(x)$ is integrable for $a \le x \le a+h$ 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) $\lim_{n\to\infty} \frac{1}{n}\left(\sin\left(\frac{y}{n}\right) + \sin\left(\frac{2y}{n}\right) + \cdots + \sin y\right)$, (d) $\lim_{n\to\infty} n^{-1}\left[(n+1)(n+2)\dots(2n)\right]^{1/n}$.

label:
ex:vp1-67e

Exercise 1.77

If the functions $f(x)$ and $g(x)$ are differentiable find expressions for the first derivative of the functions

$$F(u) = \int_0^u dx\, \frac{f(x)}{\sqrt{u^2 - x^2}} \quad \text{and} \quad G(u) = \int_0^u dx\, \frac{g(x)}{(u-x)^a} \quad \text{where} \quad 0 < a < 1.$$

Note: this is a fairly difficult problem. The formula 1.47 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 $F(x)$ and $G(x)$ are differentiable. Thus we expect an equivalent to formula 1.47 to exist.

1.5 Solutions for chapter 1

Solution for Exercise 1.1 label: ex:vp1-nota01

Take the minimum of the four distances of the point from each side and draw a circle of smaller radius around this point. The interior of such circles are open sets.

Solution for Exercise 1.2 label: ex:vp1-ord01

If $f(x) = O(x^2)$ as $x \to 0$ then $f(x) < C|x^2| < C|x|$ and hence $f(x) = O(x)$.

Solution for Exercise 1.3 label: ex:vp1-ord02

(a) $x\sqrt{1+x} = x + \frac{1}{2}x^2 + \cdots = O(x)$.

(b) $x/(1+x) = x(1 - x + x^2 + \cdots) = O(x)$.

(c) $x^{3/2}/[1 - \exp(-x)] = x/[1 - (1 - x + x^2/2 + \cdots)] = x^{1/2}/[1 + O(x)] = O(x^{1/2})$.

Solution for Exercise 1.4 label: ex:vp1-ord03

(a) $x/(x-1) = (1 - 1/x)^{-1} = 1 + 1/x + \cdots = O(1)$.

(b) $\sqrt{4x^2 + x} - 2x = 2x\sqrt{1 + 1/4x} - 2x = 2x(1 + 1/8x + O(x^{-2})) - 2x = O(1)$.

(c) $(x+b)^a - x^a = x^a(1 + b/x)^a - x^a = x^a(1 + ab/x + \cdots) - x^a = O(x^{a-1})$.

Solution for Exercise 1.5 label: ex:vp1-ord04

(a) Since $x/\sqrt{x^2+y^2} \le 1$, $y/\sqrt{x^2+y^2} \le 1$ it follows that $f_k = O(f)$, $k = 1, 2$.

(b) Put $x = r\cos\theta$, $y = r\sin\theta$ so

$$\frac{\phi}{f} = a\cos^2\theta + b\sin\theta\cos\theta + c\sin^2\theta < |a| + |b| + |c| = O(1),$$

and $\phi = O(f)$. If $y = kx$ with $2kc = -b \pm \sqrt{b^2 - 4ac}$ then $\phi = 0$.

Solution for Exercise 1.6 label: ex:vp1-lim01

Since $f(0) = 1$, we must have $A = 0$ and $B = 1$. Since $f(x)$ is finite as $x \to \infty$, $D = 0$. At $x = a$, $f(x)$ is continuous and hence $a + 1 = C/a^2$.

Solution for Exercise 1.7 label: ex:vp1-lim02

(a) $\displaystyle\lim_{x\to 0}\frac{\sin ax}{x} = a\lim_{x\to 0}\frac{\sin ax}{ax} = a,$ (b) $\displaystyle\lim_{x\to 0}\frac{\tan ax}{x} = \lim_{x\to 0}\frac{\sin ax}{x}\frac{1}{\cos ax} = a$

(c) $\displaystyle\lim_{x\to 0}\frac{\sin ax}{\sin bx} = \lim_{x\to 0}\frac{\sin ax}{x}\frac{x}{\sin bx} = \frac{a}{b},$ (d) $\displaystyle\lim_{x\to 0}\frac{3x+4}{4x+2} = \lim_{x\to 0}(3x+4)\lim_{x\to 0}\frac{1}{4x+2} = 2.$

For part (e) Take the logarithm then if E is the limit, $\ln E = \displaystyle\lim_{w\to\infty} w\ln\left(1 + \frac{z}{w}\right) = z$ so $E = e^z$.

Solution for Exercise 1.8 label: ex:vp1-lim03

In all cases we use the polar coordinates $x = r\cos\theta$, $y = r\sin\theta$.

(a) $f = \sin 2\theta$, which is independent of r, so the value of the function at the origin depends upon the direction of approach, that is θ, so f is not continuous.

(b) $f = 1/\cos 2\theta$; the same remark as in part (a) applies and f is not continuous

(c) $f = r\cos\theta\sin 2\theta$, so $f \to 0$ as $r \to 0$ independent of θ and the function is continuous.

label: ex:vp1-mon01

Solution for Exercise 1.9

(a) Since $y' = 3(a^2 - x^2)$, y is strictly increasing on $(-a, a)$. At $x = \pm a$, $y = \pm 2a^3$.

(b) With $x = 2a\sin\phi$, $|\phi| < \sin^{-1}(1/2) = \pi/6$,

$$y = 6a^3 \sin\phi - 2a^3\left(3\sin\phi - \sin 3\phi\right) = 2a^3 \sin 3\phi.$$

Hence

$$\phi = \frac{1}{3}\sin^{-1}\left(\frac{y}{2a^3}\right) \quad \text{and} \quad x(y) = 2a\sin\left(\frac{1}{3}\sin^{-1}\left(\frac{y}{2a^3}\right)\right), \quad |y| < 2a^3.$$

(c) For $x > a$, $y(x)$ is strictly decreasing and for $x > 2a$, $y < -2a^3$. Set $x = 2a\cosh\phi$ and the equation becomes

$$y = 6a^3\cosh\phi - 2a^3\left(3\cosh\phi + \cosh 3\phi\right) = -2a^3\cosh 3\phi$$

giving

$$x(y) = 2a\cosh\left(\frac{1}{3}\cosh^{-1}\left(-\frac{y}{2a^3}\right)\right), \quad y < -2a^3.$$

label: ex:vp1-der01

Solution for Exercise 1.10

(a) Use the product and chain rule,

$$\frac{d}{dx}\left(\sqrt{a-x}\,\sqrt{b+x}\right) = \frac{\sqrt{a-x}}{2\sqrt{b+x}} - \frac{\sqrt{b+x}}{2\sqrt{a-x}} = \frac{a-b-2x}{2\sqrt{(b+x)(a-x)}}.$$

Alternatively, if $y = \sqrt{a-x}\sqrt{b+x}$, then

$$\ln y = \frac{1}{2}\ln(a-x) + \frac{1}{2}\ln(b+x) \text{ giving } \frac{1}{y}\frac{dy}{dx} = \frac{1}{2(b+x)} - \frac{1}{2(a-x)} = \frac{a-b-2x}{2(b+x)(a-x)}$$

which, on simplification, gives the same result.

(b) Define

$$y^2 = a\sin^2 x + b\cos^2 x \quad \text{to give} \quad 2y\frac{dy}{dx} = 2(a-b)\sin x\cos x \quad \text{or} \quad \frac{dy}{dx} = \frac{(a-b)\sin 2x}{2\sqrt{a\sin^2 x + b\cos^2 x}}$$

which can also be expressed in the form

$$\frac{dy}{dx} = \frac{(a-b)\sin 2x}{\sqrt{2(a+b) + 2(b-a)\cos 2x}}.$$

(c) Use the chain and product rule

$$\frac{d}{dx}\left(\cos\left(x^3\right)\cos x\right) = -3x^2 \sin\left(x^3\right)\cos x - \cos\left(x^3\right)\sin x.$$

(d) If $y = x^x = e^{x \ln x}$, putting $u = x\ln x$ the chain rule gives $\dfrac{dy}{dx} = e^u \dfrac{du}{dx} = (1+\ln x)x^x$.

Solution for Exercise 1.11 label: ex:vp1-der02

Differentation with respect to y gives $1 = \dfrac{dx}{dy}\cos x$, but $\cos x = \sqrt{1-\sin^2 x} = \sqrt{1-y^2}$, hence the result.

Solution for Exercise 1.12 label: ex:vp1-der03

(a) Since $y = f(g(y))$ differentiation with respect to y gives

$$1 = \frac{d}{dy}\Big(f(g(y))\Big) = \frac{df}{dg}\frac{dg}{dy} = f'(g)g'(y).$$

Since $\dfrac{dy}{dx} = f'(x)$ and $\dfrac{dx}{dy} = g'(y)$, the result follows.

(b) Differentiate again with respect to y

$$\frac{d^2x}{dy^2} = \frac{d}{dy}\left(\frac{dy}{dx}\right)^{-1} = \frac{d}{dx}\left(\frac{dy}{dx}\right)^{-1}\frac{dx}{dy} = -\frac{d^2y}{dx^2}\left(\frac{dy}{dx}\right)^{-2}\frac{dx}{dy} = -\frac{d^2y}{dx^2}\left(\frac{dy}{dx}\right)^{-3}.$$

Solution for Exercise 1.13 label: ex:vp1-der03b

Use the chain rule with $u = -x$, so, if $f(x)$ is even, $f(u) = f(x)$ and differentiate with respect to u, $f'(u) = f'(x)\dfrac{dx}{du} = -f'(x)$, that is $f'(-x) = -f'(x)$ and $f'(x)$ is an odd function. Examples of even functions and their derivatives, in brackets are $\cos x$ $(-\sin x)$, e^{-x^2} $(-2xe^{-x^2})$. A similar analysis applies to odd functions.

Solution for Exercise 1.14 label: ex:vp1-der04a

We have

$$\frac{1}{f(x+h)} - \frac{1}{f(x)} = -\left(\frac{f(x+h)-f(x)}{f(x+h)f(x)}\right)$$

so that

$$\lim_{h\to 0}\frac{1}{h}\left(\frac{1}{f(x+h)} - \frac{1}{f(x)}\right) = \lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\right)\frac{1}{f(x+h)f(x)} = -\frac{f'(x)}{f(x)^2}$$

The product rule is proved by writing

$$f(x+h)g(x+h) - f(x)g(x) = \big[f(x+h)-f(x)\big]\,g(x+h) + f(x)\,\big[g(x+h)-g(x)\big]$$

dividing by h and taking the limit $h \to 0$.

Solution for Exercise 1.15 label: ex:vp1-der05

The first two results follow directly by applying the product rule. Thus

$$h' = f'g + fg' \quad \Longrightarrow \quad h'' = (f''g + f'g') + (f'g' + fg'').$$

The expression for $h^{(3)}$ follows similarly. Since $\binom{2}{0} = \binom{2}{2} = 1$ and $\binom{2}{1} = 2$ the general result quoted is therefore true for $n = 1$ and 2. Suppose it to be true for n; a further differentiation gives

$$\begin{aligned} h^{(n+1)} &= \sum_{k=0}^{n} \binom{n}{k} \left(f^{(n-k+1)} g^{(k)} + f^{(n-k)} g^{(k+1)} \right) \\ &= \sum_{k=0}^{n} \binom{n}{k} f^{(n-k+1)} g^{(k)} + \sum_{s=0}^{n+1} \binom{n}{s-1} f^{(n+1-s)} g^{(s)} \quad \text{(with } s = k+1 \text{ in second sum)} \\ &= \binom{n+1}{0} f^{(n)} g^{(0)} + \binom{n}{n} f^{(0)} g^{(n)} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1} \right] f^{(n-k+1)} g^{(k)}. \end{aligned}$$

But, for all m, $\binom{m}{0} = \binom{m}{m} = 1$ and

$$\begin{aligned} \binom{n}{k} + \binom{n}{k-1} &= \frac{n!}{k!\,(n-k)!} + \frac{n!}{(k-1)!\,(n+1-k)!} = \frac{n!}{k!\,(n-k)!} + \frac{(n+1)!}{k!\,(n+1-k)!} \frac{k}{n+1} \\ &= \frac{(n+1)!}{k!\,(n+1-k)!} \frac{n+1-k}{n+1} + \frac{(n+1)!}{k!\,(n+1-k)!} \frac{k}{n+1} = \binom{n+1}{k}. \end{aligned}$$

Hence the $(n + 1)$ derivative can be written as

$$\begin{aligned} h^{(n+1)} &= \binom{n+1}{0} f^{(n+1)} g + \binom{n+1}{n+1} f g^{(n+1)} + \sum_{k=1}^{n} \binom{n+1}{k} f^{(n+1-k)} g^{(k)}, \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(n+1-k)} g^{(k)}. \end{aligned}$$

Thus, if the formula is true for n, it is true for $n + 1$: it is true for $n = 2$ and hence is true for all n.

label: ex:vp1-der06

Solution for Exercise 1.16

The chain rule with $u = f(x)$ gives $\dfrac{d}{dx} \ln u = \dfrac{du}{dx}\dfrac{1}{u} = \dfrac{f'(x)}{f(x)}$. Take the logarithm of $p(x)$ to obtain

$$\ln p = \sum_{k=1}^{n} \ln f_k(x) \quad \text{and hence} \quad \frac{p'}{p} = \sum_{k=1}^{n} \frac{f_k'(x)}{f_k(x)},$$

which is valid provided none of the $f_k(x)$ are zero, that is $p(x) \neq 0$.

label: ex:vp1-der07

Solution for Exercise 1.17

Expanding the determinant gives

$$D(x) = \begin{vmatrix} f(x) & g(x) \\ \phi(x) & \psi(x) \end{vmatrix} = f\psi - g\phi \quad \text{giving} \quad D' = (f'\psi - g'\phi) + (f\psi' - g\phi')$$

which can be put in the form quoted.

The third-order determinant, with each element a function of x,

$$D(x) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

can be written as a sum of three second-order determinants,

$$D(x) = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

Now differentiate this expression using the rule just obtained for second-order determinants; then recombine the 9 terms into a third-order determinant, to obtain

$$D'(x) = \begin{vmatrix} a' & b' & c' \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} a & b & c \\ d' & e' & f' \\ g & h & i \end{vmatrix} + \begin{vmatrix} a & b & c \\ d & e & f \\ g' & h' & i' \end{vmatrix}.$$

Solution for Exercise 1.18 label: ex:vp1-mean01

We have $F(x) = f(g(x))$ and so

$$\begin{aligned} \frac{F(x+h)-F(x)}{h} &= \frac{f(g(x+h))-f(g(x))}{h} = \frac{1}{h}\left[f\left(g(x)+hg'(x+\theta h)\right) - f\left(g(x)\right)\right] \\ &= \frac{1}{h}\left[hf'\left(g(x)+h\phi g'\right)\, g'(x+\theta h)\right]. \end{aligned}$$

where $0<\theta,\phi<1$. This gives the required result on taking the limit $h\to 0$.

Solution for Exercise 1.19 label: ex:vp1-mean02

(a) $\dfrac{1}{x}\displaystyle\int_0^x dt\,\sqrt{4+4t^3} = \sqrt{4+3(\theta x)^2}$ for $0<\theta<1$. Hence the limit is 2.

(b) $\dfrac{1}{(x-1)^3}\displaystyle\int_0^x dt\,\ln\left(3t-3t^2+t^3\right) = \dfrac{1}{z^3}\displaystyle\int_0^z ds\,\ln(1+s^3)$ where $z=x-1$ and $s=t-1$. the Mean Value theorem gives the second integral as $z^{-2}\ln(1+(z\theta)^3)$, $0<\theta<1$ and this is zero in the limit $z\to 0$.

Solution for Exercise 1.20 label: ex:vp1-part01

(a) We have $\dfrac{\partial u}{\partial x} = 2x\sin(\ln y)$, $\quad \dfrac{\partial u}{\partial y} = \dfrac{x^2}{y}\cos(\ln y)$.

(b) Differentiating r^2 with respect to x and y gives, respectively

$$2r\frac{\partial r}{\partial x} = 2x \quad\text{and}\quad 2r\frac{\partial r}{\partial y} = 2y,$$

hence the result. Alternatively, put $r=\sqrt{x^2+y^2}$ to obtain

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{r} \quad\text{and}\quad \frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2+y^2}} = \frac{y}{r}.$$

Solution for Exercise 1.21

label: ex:vp1-part02

Differentiating with respect to x and y gives

$$\frac{\partial\Phi}{\partial x} = -\frac{2x}{y}\exp\left(-\frac{x^2}{y}\right) = -\frac{2x}{y}\Phi \quad \text{and} \quad \frac{\partial\Phi}{\partial y} = \frac{x^2}{y^2}\exp\left(-\frac{x^2}{y}\right) = \frac{x^2}{y^2}\Phi.$$

A second differentation of the first result with respect to x gives

$$\frac{\partial^2\Phi}{\partial x^2} = -\frac{2}{y}\Phi - \frac{2x}{y}\frac{\partial\Phi}{\partial x} = -\frac{2}{y}\Phi + 4\frac{x^2}{y^2}\Phi = 4\frac{\partial\Phi}{\partial y} - \frac{2}{y}\Phi.$$

label: ex:vp1-part03

Solution for Exercise 1.22

The derivatives u_x and u_y are found in exercise 1.20(a); differentating u_y again with respect to y gives $u_{yy} = -\frac{x^2}{y^2}\left(\cos(\ln y) + \sin(\ln y)\right)$. These expressions for u_x, u_y and u_{yy} satisfy the given equation.

label: ex:vp1-part05

Solution for Exercise 1.23

In this example $f_x = y$, $f_y = x - 2yt$, $f_t = -y^2$, $dx/dt = 2t$ and $dy/dt = 3t^2$. Hence equation 1.21 becomes

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

Alternatively, express f in terms of t,

$$f(t) = t^5 - t^7 \quad \text{so} \quad \frac{df}{dt} = t^4\left(5 - 7t^2\right).$$

Using the first expresion for df/dt we have

$$\frac{\partial}{\partial y}\left(\frac{df}{dt}\right) = \frac{\partial}{\partial y}\left(x\frac{dy}{dt} + y\frac{dx}{dt} - y^2 - 2ty\frac{dy}{dt}\right) = \frac{dx}{dt} - 2y - 2t\frac{dy}{dt} = 2t\left(1 - 4t^2\right).$$

Alternatively,

$$\frac{d}{dt}\left(\frac{\partial f}{\partial y}\right) = \frac{d}{dt}(x - 2ty) = \frac{dx}{dt} - 2y - 2t\frac{dy}{dt}.$$

label: ex:vp1-part05a

Solution for Exercise 1.24

If $F = \sqrt{1 + x_1x_2}$ then the chain rule gives

$$\frac{dF}{dt} = \frac{\partial F}{\partial x_1}x_1' + \frac{\partial F}{\partial x_2}x_2' = \frac{x_1x_2' + x_1'x_2}{2\sqrt{1 + x_1x_2}}.$$

Alternatively, set $u = x_1x_2$, so $\frac{dF}{dt} = \frac{1}{2\sqrt{1+u}}\frac{du}{dt}$, which is a simpler method of deriving the same result.

Differentiate this expression with respect to x_1, using the product rule,

$$\begin{aligned} \frac{\partial}{\partial x_1}\left(\frac{dF}{dt}\right) &= (x_1x_2' + x_1'x_2)\frac{\partial}{\partial x_1}\left(\frac{1}{2\sqrt{1 + x_1x_2}}\right) + \frac{1}{2\sqrt{1 + x_1x_2}}\frac{\partial}{\partial x_1}(x_1x_2' + x_1'x_2) \\ &= -\frac{1}{4}(x_1x_2' + x_1'x_2)\frac{x_2}{(1 + x_1x_2)^{3/2}} + \frac{x_1'}{2\sqrt{1 + x_1x_2}}. \end{aligned}$$

Also $\dfrac{\partial F}{\partial x_1} = \dfrac{x_2}{2\sqrt{1+x_1x_2}}$, and the chain rule gives

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

as before.

Solution for Exercise 1.25 label: ex:vp1-part06

(a) We have using equation 1.19

$$\frac{d}{d\lambda}f(\lambda x, \lambda y) = \frac{\partial f}{\partial u}\frac{du}{d\lambda} + \frac{\partial f}{\partial v}\frac{dv}{d\lambda} \quad \text{where} \quad u = \lambda x \quad \text{and} \quad v = \lambda y.$$

Now substitute $f(\lambda x, \lambda y) = \lambda^p f(x,y)$ into the left hand side of give

$$\frac{d}{d\lambda}f(\lambda x, \lambda y) = p\lambda^{p-1}f(x,y)$$

and set $\lambda = 1$ to obtain the result.

(b) Differentiate both sides of the relation

$$\lambda^p f(x_1, x_2, \cdots, x_n) = f(\lambda x_1, \lambda x_2, \cdots, \lambda x_n)$$

with respect to λ to obtain

$$p\lambda^{p-1}f(\mathbf{x}) = \sum_{k=1}^{n} f_{x_k}(\lambda\mathbf{x})\frac{\partial \lambda x_k}{\partial \lambda} = \sum_{k=1}^{n} x_k f_{x_k}(\lambda\mathbf{x}),$$

and set $\lambda = 1$.

(c) Differentiate both sides of the relation with respect to x_k to obtain

$$\lambda^p f_{x_k}(\mathbf{x}) = \frac{\partial}{\partial x_k}f(\lambda\mathbf{x}) = \lambda\frac{\partial}{\partial \lambda x_k}f(\lambda\mathbf{x}) = \lambda f_{x_k}(\lambda\mathbf{x})$$

which proves the result.

Solution for Exercise 1.26 label: ex:vp1-imp03

We have $f_x = -g'(x)$ and $f_y = 1$ and equations 1.22 and 1.23 give $\dfrac{dy}{dx} = g'(x)$ and $\dfrac{dx}{dy} = -1/g'(x)$, hence the result.

Solution for Exercise 1.27 label: ex:vp1-imp01

Here $f(x,y) = \ln(x^2+y^2) - 2\tan^{-1}(y/x)$ giving

$$\begin{aligned} f_x &= \frac{2x}{x^2+y^2} + \frac{2}{(1+y^2/x^2)}\frac{y}{x^2} = \frac{2(x+y)}{x^2+y^2}, \\ f_y &= \frac{2y}{x^2+y^2} - \frac{2}{x}\frac{1}{(1+y^2/x^2)} = \frac{2(y-x)}{x^2+y^2}. \end{aligned}$$

Hence $\dfrac{dy}{dx} = -\dfrac{f_x}{f_y} = \dfrac{x+y}{x-y}$.

Solution for Exercise 1.28 label: ex:vp1-imp02

Assuming $y(0)$ is finite, putting $x=0$ in the equation gives $y(0)=0$. If $f = x - y + \sin(xy)$ then $\dfrac{dy}{dx} = -\dfrac{f_x}{f_y} = \dfrac{1+y\cos(xy)}{1-x\cos(xy)}$ and hence $y'(0)=1$. Rewrite the expression for $y'(x)$ in the form $(x\cos u - 1)\,y'(x) + 1 + y\cos u = 0, \quad u = xy$ and differentiate to obtain

$$(x\cos u - 1)\,y''(x) + (\cos u - xu'\sin u)\,y'(x) + y'(x)\cos u - yu'\sin u = 0.$$

But $u' = xy' + y$, which is zero at $x=0$. Hence at $x=0$ this equation becomes $-y''(0) + 2y'(0) = 0$ and hence $y''(0) = 2$.

label: ex:vp1-imp04 **Solution for Exercise 1.29**

If $y = xv(x)$ the equation for v is

$$x\frac{dv}{dx} = -\frac{a^2+v^2}{v+1} \quad \text{or} \quad \int dv\,\frac{v+1}{v^2+a^2} = -\int\frac{dx}{x}.$$

Integration and substituting for $v = y/x$ then gives

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

where A is a constant. Since $y(1)=\alpha$ we obtain the given expression for A.

label: ex:vp1-imp05 **Solution for Exercise 1.30**

The Jacobian determinant for the functions $f_1(r,\theta) = r\cos\theta$ and $f_2(r,\theta) = r\sin\theta$ is

$$J = \begin{vmatrix} \dfrac{\partial f_1}{\partial r} & \dfrac{\partial f_1}{\partial \theta} \\ \dfrac{\partial f_2}{\partial r} & \dfrac{\partial f_2}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$$

Hence, provided $r \neq 0$, $J > 0$ and the equations may be inverted. Squaring and adding gives $r = \sqrt{x^2+y^2}$; division gives $\theta = tan^{-1}(y/x)$.

label: ex:vp1-tay01 **Solution for Exercise 1.31**

We have $f'(x) = ie^{ix}$ and $f''(x) = i^2e^{ix}$. Assuming $f^{(n)}(x) = i^ne^{ix}$ differentiating and using induction, we see that the result holds for all n. Equation 1.28 for the Taylor series, with $a=0$ then gives

$$f(x) = \sum_{k=0}^{\infty}\frac{(ix)^k}{k!}.$$

In this example $a_n = i^n/n!$, so $|a_n/a_{n+1}| = n+1 \to \infty$ as $n \to \infty$ so the radius of convergence is infinite.

Since $i^{2k} = (-1)^k$ and $i^{2k+1} = i(-1)^k$ we can write the series as the form

$$e^{ix} = \sum_{p=0}^{\infty}\frac{(ix)^{2p}}{(2p)!} + \sum_{q=0}^{\infty}\frac{(ix)^{2q+1}}{(2q+1)!} = \sum_{p=0}^{\infty}\frac{(-x^2)^p}{(2p)!} + i\sum_{q=0}^{\infty}\frac{(-1)^q(x)^{2q+1}}{(2q+1)!}.$$

But $e^{ix} = \cos x + i \sin x$, so equating real and imaginary parts gives the quoted series.

Solution for Exercise 1.32

label: ex:vp1-tay02

If $f(x) = (1+x)^a$ then $f'(x) = a(1+x)^{a-1}$, $f''(x) = a(a-1)(1+x)^{a-2}$ and

$$f^{(k)}(x) = a(a-1)(a-2)\cdots(a-k+1)(1+x)^{a-k} \quad \text{for all } k \text{ provided } k \text{ is not an integer.}$$

Thus the Taylor series about the origin becomes

$$(1+x)^a = \sum_{k=0}^{\infty} \frac{a(a-1)(a-2)\cdots(a-k+1)}{k!} x^k.$$

If a is an integer, $a = n$, this series terminates at $k = n$, to give the usual binomial expansion of $(1+x)^n$. In this example, when a is not an integer, we see that $|a_{k+1}/a_k| = |(k+1)/(a-k+2)| \to 1$ as $k \to \infty$, so the radius of convergence is unity.

Solution for Exercise 1.33

label: ex:vp1-tay03

Since $f = \sin x/\cos x$, is an odd function only odd powers occur in the Taylor expansion: we have

$$f'(x) = 1 + \frac{\sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}, \quad f''(x) = \frac{2\sin x}{\cos^3 x} \quad \text{and} \quad f^{(3)}(x) = \frac{2}{\cos^2 x} + \frac{6\sin^2 x}{\cos^4 x},$$

and $f(0) = f''(0) = 0$ (as expected) and $f'(0) = 1$ and $f^{(3)}(0) = 2$ giving the required Taylor series.

Solution for Exercise 1.34

label: ex:vp1-tay04

(a) For the first part use the solution of exercise 1.32, with $a = -1$ so $a(a-1)\cdots(a-k+1) = (-1)^k k!$, giving the quoted series. Then

$$\begin{aligned} \ln(1+x) &= \int_0^x dt\, \frac{1}{1+t} = \int_0^x dt\, \left(1 - t + t^2 + \cdots + (-1)^n t^n + \cdots\right) \\ &= x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{(-1)^n x^{n+1}}{n+1} \cdots . \end{aligned}$$

(b) The series for $(1+t)^{-1}$ is valid for $|t| < 1$, so for $|x| < 1$ the integral and sum may be interchanged.

(c) Put $x \to -x$ and subtract this from the original series.

Solution for Exercise 1.35

label: ex:vp1-tay05

The series is obtained from the solution of the previous exercise by replacing t with t^2. Then

$$\tan^{-1} x = \int_0^x dt\, \frac{1}{1+t^2} = \int_0^x dt\, \sum_{k=0}^{\infty} (-1)^k t^{2k} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1}.$$

Solution for Exercise 1.36

label: ex:vp1-tay06

Use the two Taylor expansions

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + O(z^5) \quad \text{and} \quad \sinh x = x\left(1 + \frac{x^2}{6}\right) + O(x^5)$$

to give

$$\begin{aligned}\ln(1+\sinh x) &= x\left(1+\frac{x^2}{6}+\cdots\right) - \frac{x^2}{2}\left(1+\frac{x^2}{6}+\cdots\right)^2 + \frac{x^3}{3} - \frac{x^4}{4} + O(x^5)\\ &= \left(x+\frac{x^3}{6}\right) - \left(\frac{x^2}{2}+\frac{x^4}{6}\right) + \frac{x^3}{3} - \frac{x^4}{4} + O(x^5)\\ &= x - \frac{x^2}{2} + \frac{x^3}{2} - \frac{5x^4}{12} + O(x^5).\end{aligned}$$

label: ex:vp1-tay07

Solution for Exercise 1.37

If $y(0)=0$, putting $x=0$ gives $y'(0)=1$. Differentiate n times using Leibniz's rule:

$$(1+x)y^{(n+1)} + ny^{(n)} = xy^{(n)} + ny^{(n-1)} + \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} y^{(k)} y^{(n-k)}.$$

With $n=1$, 2, 3, and 4 this gives

$n=1 \quad (1+x)y^{(2)} + y^{(1)} = xy' + y + 2yy'$

$n=2 \quad (1+x)y^{(3)} + 2y^{(2)} = xy'' + 2y' + 2y'^2 + 2yy''$

$n=3 \quad (1+x)y^{(4)} + 3y^{(3)} = xy^{(3)} + 3y^{(2)} + 2y^{(3)}y + 6y^{(2)}y^{(1)}$

$n=4 \quad (1+x)y^{(5)} + 4y^{(4)} = xy^{(4)} + 4y^{(3)} + 2y^{(4)}y + 8y^{(3)}y^{(1)} + 6\left(y^{(2)}\right)^2$

Since $y(0)=0$ and $y^{(1)}(0)=1$ (from the original equation) these equations give $y^{(2)}(0)=-1$, $y^{(3)}(0)=6$, $y^{(4)}(0)=-27$ and $y^{(5)}(0)=186$ and hence

$$y = x - \frac{1}{2}x^2 + x^3 - \frac{9}{8}x^4 + \frac{31}{20}x^5 + O(x^6).$$

An alternative method is to assume the expansion $y = x + \sum_{k=2}^{5} a_k x^k$, which automatically satisfies the conditions $y(0)=0$ and $y'(0)=1$, to substitute this into the differential equation, collect the powers of x^k, $k=2,3,\cdots,5$, and equate their coefficients to zero to obtain equations for the constants a_k.

label: ex:vp1-tays01

Solution for Exercise 1.38

(a) The required derivatives are

$f_x = \cos x \sin y$ giving $f_x(0,0)=0$, $\quad f_y = \sin x \cos y$ giving $f_y(0,0)=0$,

$f_{xx} = -\sin x \sin y$ giving $f_{xx}(0,0)=0$, $\quad f_{xy} = \cos x \cos y$ giving $f_{xy}(0,0)=1$,

$f_{yy} = -\sin x \sin y$ giving $f_{yy}(0,0)=0$,

and hence, to this order, $\sin x \sin y = xy$, as might be expected from the Taylor series for each component of the product.

(b) Put $u(x,y) = x + e^{-y} - 1$, so $u(0,0)=0$ and use the chain and product rule to compute the derivatives, $f_x = u_x \cos u$, $f_y = u_y \cos u$, $f_{xx} = u_{xx}\cos u - u_x^2 \sin u$,

$f_{yy} = u_{yy}\cos u - u_y^2 \sin u$, $f_{xy} = u_{xy}\cos u - u_y u_x \sin u$. Since $u_x = 1$, $u_{xx} = u_{xy} = 0$, $u_y = -e^{-y}$ and $u_{yy} = e^{-y}$ we obtain,

$$f_x(0,0) = 1, \quad f_y(0,0) = -1, \quad f_{xx}(0,0) = 0, \quad f_{xy}(0,0) = 0, \quad f_{yy}(0,0) = 1,$$

and hence $f = (x - y) + \frac{1}{2}y^2 + \cdots$.

Solution for Exercise 1.39

label: ex:vp1-hop01

(a) $\lim_{x\to a} \frac{\cosh x - \cosh a}{\sinh x - \sinh a} = \lim_{x\to a} \frac{\sinh x}{\cosh x} = \tanh a$.

(b)

$$\lim_{x\to 0} \frac{\sin x - x}{x\cos x - x} = \lim_{x\to 0} \frac{\cos x - 1}{\cos x - x\sin x - 1} = \lim_{x\to 0} \frac{\sin x}{x\cos x + 2\sin x} = \lim_{x\to 0} \frac{\cos x}{3\cos x - x\sin x} = \frac{1}{3},$$

(c) $\lim_{x\to 0} \frac{3^x - 3^{-x}}{2^x - 2^{-x}} = \lim_{x\to 0} \frac{e^{x\ln 3} - e^{-x\ln 3}}{e^{x\ln 2} - e^{-x\ln 2}} = \lim_{x\to 0} \frac{\ln 3}{\ln 2}\frac{e^{x\ln 3} + e^{-x\ln 3}}{e^{x\ln 2} + e^{-x\ln 2}} = \frac{\ln 3}{\ln 2}$.

Solution for Exercise 1.40

label: ex:vp1-hop02

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

(b) Put $F(x) = 1/f(x)$ and $G(x) = 1/g(x)$ so $F(a) = G(a) = 0$, and

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{G(x)}{F(x)} = \lim_{x\to a} \frac{g'(x)}{f'(x)}\frac{f(x)^2}{g(x)^2} = \left(\lim_{x\to a} \frac{g'(x)}{f'(x)}\right)\left(\lim_{x\to a} \frac{f(x)}{g(x)}\right)^2.$$

Hence, provided all limits exist, $\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$.

Solution for Exercise 1.41

label: ex:vp1-int01a

(a) $\int_{-a}^{a} dx\, f(x) = \int_{-a}^{0} dx\, f(x) + \int_{0}^{a} dx\, f(x)$, put $x = -u$ in the first integral and use the fact that $f(-u) = -f(u)$ to show that the two integrals have the same magnitude but opposite signs.

(b) Split the integral in the same manner as in part (a), but since $f(-u) = f(u)$ the two integrals are equal.

Solution for Exercise 1.42

label: ex:vp1-int01b

Assuming $\lambda > 0$, put $y = \lambda x$ in the integral, which becomes $I(\lambda) = \int_0^\infty dy\, \frac{\sin y}{y}$.

If $\mu > 0$, put $\lambda = -\mu$ to obtain $I(-\mu) = -\int_0^\infty dx\, \frac{\sin \mu x}{x} = -I(\mu)$.

Solution for Exercise 1.43

label: ex:vp1-int01

(a) $\int dx\, 1 \times \ln x = x \ln x - \int dx = -x(1-\ln x)$.

(b) $\int dx\, \dfrac{x}{\cos^2 x} = x\tan x - \int dx\, \tan x$ but $\int dx\, \tan x = \int dx\, \dfrac{\sin x}{\cos x} = -\ln|\cos x|$.
Hence $\int dx\, \dfrac{x}{\cos^2 x} = x\tan x + \ln|\cos x|$.

(c) $\int dx\, x\ln x = \dfrac{1}{2}x^2 \ln x - \dfrac{1}{2}\int dx\, x^2 \dfrac{1}{x} = \dfrac{1}{2}x^2\ln x - \dfrac{1}{4}x^2$.

(d) $\int dx\, x\sin x = -x\cos x + \int dx\, \cos x = \sin x - x\cos x$.

label: ex:vp1-int02

Solution for Exercise 1.44

(a) Put $\cos x = u$ to obtain

$$\int_{1/\sqrt{2}}^{1} du\, \ln u = \left[-u(1-\ln u)\right]_{1/\sqrt{2}}^{1} = \frac{1}{2\sqrt{2}}\ln 2 + \frac{1}{\sqrt{2}} - 1.$$

(b)

$$\begin{aligned}\int_0^{\pi/4} dx\, x\tan^2 x &= \int_0^{\pi/4} dx\, \left(\frac{x}{\cos^2 x} - x\right) = \left[x\tan x + \ln\cos x - \frac{1}{2}x^2\right]_0^{\pi/4} \\ &= \frac{\pi}{4} - \frac{1}{2}\ln 2 - \frac{\pi^2}{32}.\end{aligned}$$

(c)

$$\int_0^1 dx\, x^2 \sin^{-1} x = \left[\frac{1}{3}x^3 \sin^{-1} x\right]_0^1 - \frac{1}{3}\int_0^1 dx\, \frac{x^3}{\sqrt{1-x^2}}.$$

But on putting $x = \sin\phi$ and using the identity $\sin 3\phi = 3\sin\phi - 4\sin^3\phi$,

$$\int_0^1 dx\, \frac{x^3}{\sqrt{1-x^2}} = \int_0^{\pi/2} d\phi\, \sin^3\phi = \frac{1}{4}\int_0^{\pi/2} d\phi\, (3\sin\phi - \sin 3\phi) = \frac{2}{3}$$

and hence $\int_0^1 dx\, x^2 \sin^{-1} x = \dfrac{\pi}{6} - \dfrac{2}{9}$.

label: ex:vp1-int03

Solution for Exercise 1.45

Integrating by parts for $n \geq 1$ gives

$$I_n = \int_0^x dt\, t^n e^{at} = \left[\frac{t^n}{a}e^{at}\right]_0^x - \frac{n}{a}I_{n-1}$$

and hence $aI_n = x^n e^{ax} - nI_{n-1}$, $n \geq 1$, with $I_0 = (e^{ax}-1)/a$.

The equations for I_k, $k = 1, 2, \cdots, n$, are

$$aI_1 = xe^{ax} - I_0, \quad aI_2 = x^2 e^{ax} - 2I_1, \quad aI_3 = x^3 e^{ax} - 3I_2, \; \cdots, aI_n = x^n e^{ax} - nI_{n-1}.$$

Multiply the kth equation by A_k and add all the equations to obtain

$$a\sum_{k=1}^{n} A_k I_k = e^{ax}\sum_{k=1}^{n} A_k x^k - \sum_{k=1}^{n} kA_k I_{k-1}.$$

Now chose the A_k such that $A_n = 1/a$ and for $k = 1, 2, \cdots, n-1$, the I_k cancel, that is

$$aA_k = -(k+1)A_{k+1}, \quad k = 1, 2, \cdots, n-1, \quad A_n = \frac{1}{a}.$$

The solution of these equations is $A_k = \dfrac{n!\,(-1)^{n-k}}{a^{n-k+1}\,k!}$ which gives the quoted expression.

Solution for Exercise 1.46 label: ex:vp1-int03a

(a) $\displaystyle\int_0^a dx\, f(x) = -\int_a^0 du\, f(a-u) = \int_0^a dx\, f(a-x).$

(b) Since $\sin(\pi/2 - \phi) = \cos\phi$ and $\cos(\pi/2-\phi) = \sin\phi$ we have

$$I = \int_0^{\pi/2} d\theta\, \frac{\sin\theta}{\sin\theta + \cos\theta} = -\int_{\pi/2}^{0} d\phi\, \frac{\cos\phi}{\cos\phi + \sin\phi} = \int_0^{\pi/2} d\phi\, \frac{\cos\phi}{\cos\phi+\sin\phi}.$$

Hence, on adding these two equivalent forms, $2I = \pi/2$.

Solution for Exercise 1.47 label: ex:vp1-int04

With $t = \tan(x/2)$ the integral becomes

$$\int_0^{\pi} dx\, \frac{1}{a + b\cos x} = \int_0^{\infty} dt\, \frac{dx}{dt}\frac{1}{a + b\left(\frac{1-t^2}{1+t^2}\right)} = 2\int_0^{\infty} dt\, \frac{1}{a+b+(a-b)t^2},$$

since $dt/dx = (1+t^2)/2$. The integral is evaluated with the further substitution $t = \sqrt{\dfrac{a+b}{a-b}}\tan z$, to give the quoted result. If $b > a$, $a + b\cos x = 0$ for $0 < x < \pi$, the integrand is singular and the integral does not exist.

Define $F(a,b) = \displaystyle\int_0^{\pi} dx\, \frac{1}{a+b\cos x} = \frac{\pi}{\sqrt{a^2-b^2}}$ then

$$\begin{aligned}
-\frac{\partial F}{\partial a} &= \int_0^{\pi} dx\, \frac{1}{(a+b\cos x)^2} = \frac{a\pi}{(a^2-b^2)^{3/2}},\\
\frac{\partial^2 F}{\partial a^2} &= 2\int_0^{\pi} dx\, \frac{1}{(a+b\cos x)^3} = \frac{\pi(2a^2+b^2)}{(a^2-b^2)^{5/2}},\\
-\frac{\partial F}{\partial b} &= \int_0^{\pi} dx\, \frac{\cos x}{(a+b\cos x)^2} = -\frac{b\pi}{(a^2-b^2)^{3/2}}.
\end{aligned}$$

For the last example define

$$G(a,b) = \int_0^{\pi} dx\, \ln(a + b\cos x) \quad \text{so} \quad \frac{\partial G}{\partial a} = \int_0^{\pi} dx\, \frac{1}{a+b\cos x} = \frac{\pi}{\sqrt{a^2-b^2}}.$$

Integrating with respect to a gives

$$G = C + \pi \int da \, \frac{1}{\sqrt{a^2 - b^2}} = C + \pi \ln(a + \sqrt{a^2 - b^2}),$$

where C is a constant. But if $b = 0$, $G = \pi \ln a$ and hence $C = -\pi \ln 2$ and we obtain

$$\int_0^\pi dx \, \ln(a + b\cos x) = \pi \ln \left(\frac{a + \sqrt{a^2 - b^2}}{2} \right).$$

label: ex:vp1-int05

Solution for Exercise 1.48

First note that the integral expression for $y(t)$ gives $y(a) = 0$. Differentiate twice with respect to t, using the formula 1.47,

$$\frac{dy}{dt} = \int_a^t dx \, f(x) \cos\omega(t-x) \quad \text{and} \quad \frac{d^2y}{dt^2} = f(t) - \omega \int_a^t dx \, f(x) \sin\omega(t-x) = f(t) - \omega^2 y(t).$$

From the first of these equations we see that $y'(a) = 0$, so the initial conditions are satisfied. the second equations gives $y''(a) = f(a)$, which is consistent with the original differential equation.

label: ex:vp1-int06

Solution for Exercise 1.49

(a) Since

$$F(u+h) = \int_0^{a(u+h)} dx \, f(x) = \int_0^{a(u)} dx \, f(x) + \int_{a(u)}^{a(u+h)} dx \, f(x)$$

we have

$$\frac{F(u+h) - F(u)}{h} = \frac{1}{h} \int_{a(u)}^{a(u+h)} dx \, f(x) = \frac{a(u+h) - a(u)}{h} f(\xi), \quad \text{where} \quad \xi \in \Big(a(u), a(u+h) \Big),$$

the last result being obtained from the integral form of the Mean Value Theorem. Taking the limit $h \to 0$ gives $F'(u) = a'(u)F(a(u))$. The same result can be derived using the Fundamental theorem of Calculus and the chain rule.

(b) We have

$$\frac{f(u+h) - f(u)}{h} = \int_a^b dx \, \frac{f(x, u+h) - f(x, u)}{h}$$

Assuming that the limit $h \to 0$ exists we obtain $f'(u) = \displaystyle\int_a^b dx \, \frac{\partial f}{\partial u}$.

label: ex:vp1-int07

Solution for Exercise 1.50

In the first case

$$\int_2^X dx \, \frac{1}{x \ln x} = \int_2^X dx \, \frac{d}{dx} \ln(\ln x) = \ln(\ln X) - \ln(\ln 2) \to \infty \text{ as } X \to \infty.$$

In the second case, integration by parts gives

$$\int_2^X dx\,\frac{1}{x(\ln x)^2} = \left[\frac{1}{\ln x}\right]_2^X + 2\int_2^X dx\,\frac{1}{x(\ln x)^2}$$

and hence

$$\int_2^X dx\,\frac{1}{x(\ln x)^2} = \frac{1}{\ln 2} - \frac{1}{\ln X} \to \frac{1}{\ln 2} \text{ as } X \to \infty.$$

Solution for Exercise 1.51 label: ex:vp1-01e

(a) Put $x = 1 + \delta$ so the ratio becomes

$$\frac{e^{a\ln(1+\delta)} - 1}{\delta} = \frac{e^{a\delta+O(\delta^2} - 1}{\delta} = a + O(\delta).$$

(b) Use the binomial expansion

$$\frac{\sqrt{1+x}-1}{1-\sqrt{1-x}} = \frac{\left(1 + x/2 + O(x^2)\right) - 1}{1 - (1 - x/2 + O(x^2))} = \frac{1+O(x)}{1+O(x)} = 1 + O(x).$$

(c) Put $x = a + \delta$ and use the binomial expansion to give

$$\frac{(a+\delta)^{1/3} - a^{1/3}}{(a+\delta)^{1/2} - a^{1/2}} = \frac{a^{1/3}\left(1 + \frac{\delta}{3a} + O(\delta^2)\right) - a^{1/3}}{a^{1/2}\left(1 + \frac{\delta}{2a} + O(\delta^2)\right) - a^{1/2}} = \frac{2}{3a^{1/6}} + O(\delta).$$

(d) Put $x = \pi/2 - \delta$, $\delta > 0$ to give $(\pi - 2x)\tan x = \dfrac{2\delta}{\tan\delta} = 2 + O(\delta)$.

(e) Put $y = x^{1/x}$, so $\ln y = (1/x)\ln x$ and $\lim_{x\to 0}\ln y = -\infty$ and $\lim_{x\to 0} x^{1/x} = 0$.

(f) We have

$$\lim_{x\to 0}\left(\frac{1+x}{1-x}\right)^{1/x} = \lim_{x\to 0}\exp\left(\frac{1}{x}\ln\left(\frac{1+x}{1-x}\right)\right) = \lim_{x\to 0}\exp\left(\frac{1}{x}2\left(x + O(x^3)\right)\right) = e^2$$

Solution for Exercise 1.52 label: ex:vp1-71e

In all cases put $z = e^x$. In the first example this gives $2y = z + 1/z$ or $z^2 - 2yz + 1 = 0$. This quadratic has the two solutions $z = e^x = y \pm \sqrt{y^2-1}$, one of which is larger than unity and the other smaller — because they are real and their product is unity. For $x > 0$, $e^x > 1$ and so

$$x = \ln z = \ln(y + \sqrt{y^2-1}).$$

In the second example the quadratic equation is $z^2 - 2yz - 1 = 0$, with solutions $z = e^x = y \pm \sqrt{y^2+1}$. Since $e^x > 0$ we choose the positive root to give

$$x = \ln(y + \sqrt{y^2+1}).$$

In the finally example we have $y = \dfrac{z^2-1}{z^2+1}$or $z = e^x = \pm\sqrt{\dfrac{1+y}{1-y}}$. The positive root gives the required solution so

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

label: ex:vp1-72e

Solution for Exercise 1.53

Since $y'(x) = x'(y)^{-1}$ a second differentiation gives

$$\frac{d^2y}{dx^2} = \frac{d}{dy}\left(\frac{1}{x'(y)}\right)\frac{dy}{dx} = -\frac{x''(y)}{x'(y)^3}$$

and since $y''(x) = -y$ this gives

$$\frac{d^2x}{dy^2} = y\left(\frac{dx}{dy}\right)^3 \quad \text{or} \quad \frac{dz}{dy} = yz^3 \quad \text{if} \quad z = \frac{dx}{dy}.$$

Integration gives $1/z^2 = -y^2 + c$, but $x'(0) = 1/y'(0) = 1$ and $y(0) = 0$, so $c = 1$ and

$$\frac{dx}{dy} = \frac{1}{\sqrt{1-y^2}}, \quad x(0) = 0,$$

where the negative square root is ignored because $x'(0) = 1$. A further integration gives $x(y) = \displaystyle\int_0^y du\, \frac{1}{\sqrt{1-u^2}}$.

The Taylor expansion of the integrand is

$$\frac{1}{\sqrt{1-u^2}} = 1 + \frac{1}{2}u^2 + \frac{3}{8}u^4 + O(u^6)$$

so integration gives

$$\sin^{-1} y = y + \frac{1}{6}y^3 + \frac{3}{40}y^5 + O(y^7).$$

More generally, we have $\dfrac{1}{\sqrt{1-u^2}} = \displaystyle\sum_{k=0}^{\infty} \frac{(2k)!\, u^{2k}}{k!^2 2^{2k}}$, $|u| < 1$, so

$$\sin^{-1} y = y\sum_{k=0}^{\infty} \frac{(2k)! y^{2k}}{k!^{\,2}\, 2^{2k}\,(2k+1)}, \quad |y| < 1.$$

label: ex:vp1-03e

Solution for Exercise 1.54

(a) Since $\ln y = g\ln f$ we have $y'/y = g'\ln f + gf'/g$ and hence

$$\frac{dy}{dx} = (fg'\ln f + gf')\, f(x)^{g(x)-1}.$$

(b) Since

$$\ln y = \frac{1}{2}\ln(p+x) - \frac{1}{2}\ln(p-x) + \frac{1}{2}\ln(q+x) - \frac{1}{2}\ln(g-x)$$

we have

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

and

$$\frac{dy}{dx} = \left(\frac{p}{p^2-x^2} + \frac{q}{q^2-x^2}\right)\sqrt{\frac{p+x}{p-x}}\sqrt{\frac{q+x}{q-x}}.$$

(c) We have

$$ny^{n-1}\frac{dy}{dx} = 1 + \frac{x}{\sqrt{1+x^2}} = \frac{y^n}{\sqrt{1+x^2}} \quad \text{therefore} \quad \frac{dy}{dx} = \frac{y}{n\sqrt{1+x^2}}.$$

Solution for Exercise 1.55 label: ex:vp1-04e

Differentiate using the chain rule,

$$\frac{dy}{dx} = \frac{a\cos u}{\sqrt{1-x^2}}, \quad u = a\sin^{-1}x$$

and

$$\frac{d^2y}{dx^2} = -\frac{a^2\sin x}{1-x^2} + \frac{ax\cos u}{(1-x^2)^{3/2}} = -\frac{a^2 y}{1-x^2} + \frac{x}{1-x^2}\frac{dy}{dx},$$

which gives the required result.

Solution for Exercise 1.56 label: ex:vp1-05e

If $x = \cos\theta$ we have

$$\frac{dy}{dx} = \frac{dy}{d\theta}\frac{d\theta}{dx} = -\frac{1}{\sqrt{1-x^2}}\frac{dy}{d\theta} \quad \text{since} \quad \frac{dx}{d\theta} = -\sin\theta = -\sqrt{1-x^2}.$$

and then

$$\frac{d^2y}{dx^2} = -\frac{1}{\sqrt{1-x^2}}\frac{d^2y}{d\theta^2}\frac{d\theta}{dx} - \frac{x}{(1-x^2)^{3/2}}\frac{dy}{d\theta}.$$

Hence the differential equation becomes

$$\frac{d^2y}{d\theta^2} + \frac{x}{\sqrt{1-x^2}}\frac{dy}{d\theta} + \lambda y = 0,$$

which gives the required result since $x/\sqrt{x^2+y^2} = \cot\theta$.

Solution for Exercise 1.57 label: ex:vp1-06e

Let $h(x) = f(g(x))$ then

$$\begin{aligned} h'(x) &= g'(x)f'(g), \\ h''(x) &= g''(x)f'(g) + g'(x)^2 f''(g), \\ h'''(x) &= g'''(x)f'(g) + 3g''(x)g'(x)f''(g) + g'(x)^3 f'''(g), \end{aligned}$$

so that

$$Sh(x) = \frac{g'''(x)f'(g) + 3g''(x)g'(x)f''(g) + g'(x)^3 f'''(g)}{g'(x)f'(g)} - \frac{3}{2}\left(\frac{g''(x)}{g'(x)} + \frac{g'(x)f''(g)}{f'(g)}\right)^2$$

On multiplying this out we see that $Sh(x) = Sg(x) + g'(x)^2 Sf(g) < 0$.

Solution for Exercise 1.58 label: ex:vp1-10e

Differentation gives

$$\begin{aligned}
\frac{\partial z}{\partial x} &= f' + g' - \frac{1}{2a^2}\cos(x+ay) + \frac{x}{2a^2}\sin(x+ay), \\
\frac{\partial^2 z}{\partial x^2} &= f'' + g'' + \frac{1}{a^2}\sin(x+ay) + \frac{x}{2a^2}\cos(x+ay), \\
\frac{\partial z}{\partial y} &= a(f' - g') + \frac{x}{2a}\sin(x+ay), \\
\frac{\partial^2 z}{\partial y^2} &= a^2(f'' + g'') + \frac{x}{2}\cos(x+ay),
\end{aligned}$$

and hence $a^2\dfrac{\partial^2 z}{\partial x^2} - \dfrac{\partial^2 z}{\partial y^2} = \sin(x+ay)$.

label: ex:vp1-11e **Solution for Exercise 1.59**

Differentiation gives

$$f_x = af - \frac{f}{x}, \quad f_y = bf - \frac{f}{y}, \quad \text{and} \quad f_z = cf - \frac{f}{z}.$$

So the partial derivatves are zero at $ax = by = cz = 1$.

label: ex:vp1-12e **Solution for Exercise 1.60**

Differentiate with respect to x,

$$2(uu_x - x)f_1 + 2uu_x f_2 - 2uu_x f_3 = 0 \quad \text{or} \quad uu_x(f_1 + f_2 + f_3) = xf_1$$

where $f_k = \partial f/\partial x_k$, $f = f(x_1, x_2, x_3)$. Similarly, differentiation with respect to y and z gives

$$uu_y(f_1 + f_2 + f_3) = yf_2 \quad \text{and} \quad uu_z(f_1 + f_2 + f_3) = zf_3.$$

Adding these three results gives the required equation.

label: ex:vp1-20e **Solution for Exercise 1.61**

Since $f_x = 2x$ and $f_y = 2y$ the implicit function theorem shows that $y(x)$ and $x(y)$ exist if $y \neq 0$ and $x \neq 0$, respectively, and then $y'(x) = -f_x/f_y = -x/y$.

If $f = 0$ then $y^2 = 1 - x^2$, hence $y = \pm\sqrt{1-x^2}$ and $y'(x) = \mp x/\sqrt{1-x^2} = -x/y$.

label: ex:vp1-21e **Solution for Exercise 1.62**

If $f = x\cos xy$, $f_y = -x^2\sin u$ and $f_x = \cos u - u\sin u$, where $u = xy$. Thus $f_y(1, \pi/2) = -1$ and $f_x(1, \pi/2) = -\pi/2$. Hence, from the implicit function theorem, $y(x)$ exists in the neighbourhood of $(1, \pi/2)$, with

$$y'(x) = -\frac{f_x}{f_y} = \frac{\cos u - u\sin u}{x^2\sin u} \quad \text{or} \quad x^2 y' = \frac{1}{\tan u} - u,$$

hence $y'(1) = -\pi/2$. Differentiating again gives

$$y''x^2 + 2xy' = -\left(\frac{1}{\sin^2 u} + 1\right)(xy' + y).$$

At $x = 1$, $y = \pi/2$, since $y'(1) = -\pi/2$, this gives $y'' = \pi$. Hence the Taylor expansion of $y(x)$ about $x = 1$ is

$$\begin{aligned} y(x) &= y(1) + (x-1)y'(1) + \frac{1}{2}(x-1)^2 y''(1) + \cdots \\ &= \frac{\pi}{2} - \frac{\pi}{2}(x-1) + \frac{\pi}{2}(x-1)^2 + \cdots . \end{aligned}$$

Solution for Exercise 1.63

label: ex:vp1-22e

Differentiate the equation $x^3+y^3-3axy = 0$ and re-arrange to give $(y^2-ax)y'+x^2-ay = 0$, which gives the relation for $y'(x)$. Hence y' is defined provided the denominator is not zero, that is $y^2 \neq ax$. The curve defined by $f(x, y) = 0$ is parallel to the x-axis if $x^2 = ay$, which substituted into the equation gives $x^3(x^3 - 2a^3) = 0$. At $x = 0$, y' is not defined; the solution at $x = a2^{1/3}$ gives the quoted result.

Solution for Exercise 1.64

label: ex:vp1-31e

From the graphs of $y = 1/x$ and $y = \tan x$, shown in figure 1.9, we see that the equation has positive roots x_k, $k = 0, 1, 2, \cdots$, and that $k\pi < x_k < (k+1/2)\pi$ and that for large k, $x_k \to k\pi$ from above.

label: f:vp1-sol31e

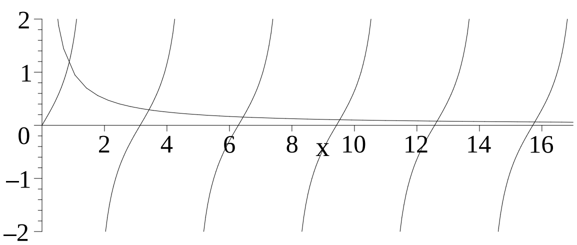

Figure 1.9 Graphs of $y = 1/x$ and $y = \tan x$.

For the nth root, put $x = n\pi + z$, and since $\sin x = (-1)^n \sin z$ and $\cos x = (-1)^n \cos z$ the equation becomes

$$(n\pi + z)\tan z = 1 \quad \text{with } z \text{ small.}$$

Put $\epsilon = 1/n\pi$ so the equation becomes $(1 + \epsilon z)\tan z = \epsilon$ and we require the Taylor expansion of $z(\epsilon)$ about $\epsilon = 0$. Putting $\epsilon = 0$ we see that $z(0) = 0$. Differentiation gives

$$(\epsilon z' + z)\tan z + \frac{1+\epsilon z}{\cos^2 z} z' = 1 \quad \text{giving} \quad z'(0) = 1,$$

and hence $x = n\pi + \dfrac{1}{n\pi}$.

Further differentiation of the same equation allows, in principle, the calculation of $z^{(n)}(0)$ for $n > 2$; however, such calculations are extremely tedious and error prone. A far easier method is now outlined.

First, rewrite the equation for z in the form

$$\tan z = \frac{\epsilon}{1+\epsilon z}$$

and observe that this equation defines a function $z(\epsilon)$, with $z(0)=0$, that is an odd function of ϵ — to see this note that $-z(-\epsilon)$ satisfies the same equation. Also, for small $|z|$ we see that to $O(\epsilon)$ the equation becomes $z=\epsilon+O(\epsilon^2)$. The power series for $z(\epsilon)$ is thus

$$z=\epsilon+z_3\epsilon^3+z_5\epsilon^5+O(\epsilon^7),$$

where z_3 and z_7 are coefficients to be found. Substitute this series in to the left hand side of the equation and use the known series for $\tan z$ to obtain

$$\begin{aligned}\tan z &= \left(\epsilon+z_3\epsilon^3+z_5\epsilon^5+\cdots\right)+\frac{\epsilon^3}{3}\left(1+z_3\epsilon^2+\cdots\right)^3+\frac{2}{15}\epsilon^5+\cdots\\ &= \epsilon+\epsilon^3\left(z_3+\frac{1}{3}\right)+\epsilon^5\left(z_5+z_3+\frac{2}{15}\right)+\cdots.\end{aligned}$$

Similarly the right hand side gives

$$\begin{aligned}\frac{\epsilon}{1+\epsilon z} &= \epsilon\left(1-\epsilon z+\epsilon^2 z^2+\cdots\right)\\ &= \epsilon-\epsilon^3+\epsilon^5\left(1-z_3\right)+\cdots.\end{aligned}$$

Equating the coefficients of the powers of ϵ on each side of the equation gives $z_3=-4/3$ and $z_5=53/15$ and hence

$$x=n\pi+\frac{1}{n\pi}-\frac{4}{3(n\pi)^2}+\frac{53}{15(n\pi)^3}+\cdots.$$

label: ex:vp1-32e

Solution for Exercise 1.65

Using the Taylor expansion of $\cos z$ the numerator becomes

$$1+a\left(1-\frac{(2x)^2}{2}+\frac{(2x)^4}{24}+\cdots\right)+b\left(1-\frac{(4x)^2}{2}+\frac{(4x)^4}{24}+\cdots\right).$$

which simplifies to

$$1+a+b-x^2(2a+8b)+x^4\left(\frac{2}{3}a+\frac{32}{3}b\right)+\cdots$$

Thus we need $a+b+1=0$, $a+4b=0$, that is $b=1/3$ and $a=-4/3$. Then the value of the function at the origin is $2a/3+32b/3=8/3$.

label: ex:vp1-33e

Solution for Exercise 1.66

There are many ways to obtain the expansions, but usually a direct use of the definition, which requires the calculation of higher derivatives, is awkward and error prone: it is usually easiest to use known results where possible. The methods outlined below are not necessarily the easiest, but just the first I thought of.

(a) Since the Taylor series of $\ln(1+z)$ is known we write, with $u = x/2$, and use the identity $\cosh 2u = 1 + 2\sinh^2 u$,

$$\begin{aligned}\ln(\cosh x) &= \ln\left(1 + 2\sinh^2 u\right)\\ &= 2\sinh^2 u - \frac{1}{2}\left(2\sinh^2 u\right)^2 + \frac{1}{3}\left(2\sinh^2 u\right)^3 + O(u^8).\end{aligned}$$

Now use

$$\sinh u = u\left(1 + \frac{u^2}{6} + \cdots\right) \quad \text{and} \quad \sinh^2 u = u^2\left(1 + \frac{u^2}{3} + \cdots\right)$$

in this expansion, to give

$$\ln(\cosh x) = 2u^2\left(1 + \frac{u^4}{12} + \cdots\right) - 2u^4 + \cdots = \frac{x^2}{2} - \frac{x^4}{12} + O(x^6)$$

(b) Similarly

$$\begin{aligned}\ln(1+\sin x) &= \sin x - \frac{1}{2}\sin^2 x + \frac{1}{3}\sin^3 x - \frac{1}{4}\sin^4 x + O(x^5).\\ \sin x &= x\left(1 - \frac{x^2}{6} + \cdots\right)\end{aligned}$$

giving

$$\begin{aligned}\ln(1+\sin x) &= x\left(1 - \frac{x^2}{6} + \cdots\right) - \frac{x^2}{2}\left(1 - \frac{x^2}{3} + \cdots\right) + \frac{x^3}{3} - \frac{x^4}{4} + O(x^5),\\ &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{12} + O(x^5).\end{aligned}$$

(c) Similarly

$$\begin{aligned}\exp(\sin x) &= 1 + \sin x + \frac{\sin^2 x}{2} + \frac{\sin^3 x}{6} + \frac{\sin^4 x}{24} + O(x^5),\\ &= 1 + x\left(1 - \frac{x^2}{6} + \cdots\right) + \frac{x^2}{2}\left(1 - \frac{x^2}{3} + \cdots\right) + \frac{x^3}{6} + \frac{x^4}{24} + O(x^5),\\ &= 1 + x + \frac{x^2}{2} - \frac{x^4}{8} + O(x^5).\end{aligned}$$

(d) Use the identity

$$\begin{aligned}\sin^2 x &= (1 - \cos 2x)/2 \quad \text{to give}\\ &= \frac{1}{2} - \frac{1}{2}\left(1 - \frac{(2x)^2}{2} + \frac{(2x)^4}{24} + O(x^6)\right) = x^2 - \frac{x^4}{3} + O(x^6).\end{aligned}$$

Solution for Exercise 1.67 label: ex:vp1-41e

The Cauchy form of the Mean Value Theorem gives $f(x+1) = f(x) + f'(x+\theta)$ with

$0 < \theta < 1$. Since $f'(x)$ is strictly increasing $f'(x) < f'(x+\theta) < f'(x+1)$ and the result follows.

Solution for Exercise 1.68

label: ex:vp1-42e

In the first case, for any $x > 0$ the Mean Value Theorem gives, for $0 < \theta < 1$,

$$f_1(x) = f_1(0) + f_1'(\theta x) = -\frac{\theta x}{1+\theta x} < 0.$$

Hence, $f_1(x) < 0$ for $x > 0$. Similary $f_2(x) > 0$ for $x > 0$.

label: ex:vp1-51e

Solution for Exercise 1.69

Using L'Hospital's rule,

$$\lim_{x\to 1} \frac{\sin \ln x}{x^5 - 7x^3 + 6} = \lim_{x\to 1} \frac{\cos \ln x}{x} \frac{1}{5x^4 - 21x^2} = -\frac{1}{16}.$$

label: ex:vp1-52e

Solution for Exercise 1.70

In the first case set $y = (\cos x)^{1/\tan^2 x}$ and consider the limit of $\ln y$,

$$\lim_{x\to 0} \frac{\ln \cos x}{\tan^2 x} = \lim_{x\to 0} \left(-\frac{\sin x}{\cos x}\frac{\cos^3 x}{2\sin x}\right) = -\frac{1}{2}, \quad \text{hence} \quad \lim_{x\to 0} y = e^{-1/2}.$$

For the second case,

$$\begin{aligned}\lim_{x\to 0} \frac{a\sin bx - b\sin ax}{x^3} &= ab \lim_{x\to 0} \frac{\cos bx - \cos ax}{3x^2} = -ab \lim_{x\to 0} \frac{b\sin bx - a\sin ax}{6x} \\ &= -\frac{ab}{6} \lim_{x\to 0} \left(b^2 \cos bx - a^2 \cos ax\right) = \frac{ab}{6}(a^2 - b^2).\end{aligned}$$

label: ex:vp1-61e

Solution for Exercise 1.71

If

$$I(a) = \int_0^\infty dx\, \frac{\tan^{-1}(ax)}{x(1+x^2)} \quad \text{then} \quad I'(a) = \int_0^\infty dx\, \frac{1}{(1+x^2)(1+a^2x^2)}.$$

Using partial fractions this becomes

$$I'(a) = \frac{1}{1-a^2} \int_0^\infty dx\, \left(\frac{1}{1+x^2} - \frac{a^2}{1+a^2x^2}\right) = \frac{1}{1-a^2}\frac{\pi}{2}(1-a) = \frac{\pi}{2}\frac{1}{1+a}.$$

Now integrate to obtain $I(a) = \frac{\pi}{2}\ln(1+a) + C$ giving $I(0) = C$. But, from the original integral $I(0) = 0$, and hence $C = 0$.

label: ex:vp1-62e

Solution for Exercise 1.72

If

$$I(z) = \int_0^{\pi/2} dx\, \frac{\ln(1+z\cos x)}{\cos x} \quad \text{then} \quad I'(z) = \int_0^{\pi/2} dx\, \frac{1}{1+z\cos x}.$$

Now use the identity $\cos x = (1-t^2)/(1+t^2)$, $t = \tan(x/2)$ to obtain

$$\begin{aligned} I'(z) &= \frac{2}{(1-z)}\int_0^1 dt\, \frac{1}{b^2+t^2}, \quad b^2 = \frac{1+z}{1-z}, \\ &= \frac{2}{\sqrt{1-z^2}}\tan^{-1}\sqrt{\frac{1-z}{1+z}}. \end{aligned}$$

But $z = \cos\pi a$ and so this becomes

$$\frac{dI}{dz} = \frac{\pi a}{\sin\pi a} \quad \text{or} \quad \frac{dI}{da} = -\pi^2 a \quad \text{and hence} \quad I(a) = C - \frac{1}{2}\pi^2 a^2.$$

But if $\cos\pi a = 0$, that is $a = 1/2$, $I = 0$. Hence $C = \pi^2/8$ and $I(a) = \dfrac{\pi^2}{8}(1-4a^2)$.

Solution for Exercise 1.73 label: ex:vp1-63e

We have

$$I = \int_0^{\pi/2} dx\, \frac{\sin x}{x}\frac{\sin(\pi/2-x)}{\pi/2-x} = \int_0^{\pi/2} dx\, \frac{\sin x\cos x}{x(\pi/2-x)} = \frac{1}{\pi}\int_0^{\pi/2} dx\, \sin 2x\left(\frac{1}{x}+\frac{1}{\pi/2-x}\right).$$

Now put $y = 2x$ to give

$$\begin{aligned} I &= \frac{1}{\pi}\int_0^{\pi} dy\, \sin y\left(\frac{1}{y}+\frac{1}{\pi-y}\right) \\ &= \frac{1}{\pi}\int_0^{\pi} dy\, \frac{\sin y}{y} + \frac{1}{\pi}\int_0^{\pi} dz\, \frac{\sin(\pi-z)}{z} = \frac{2}{\pi}\int_0^{\pi} dy\, \frac{\sin y}{y}, \quad (z = \pi - y), \end{aligned}$$

since $\sin(\pi - z) = \sin z$.

Solution for Exercise 1.74 label: ex:vp1-64e

Put $x = 1/z$ and $s = 1/t$ to obtain

$$\tan^{-1}(1/z) = \int_0^{1/z} dt\, \frac{1}{1+t^2} = \int_z^{\infty} ds\, \frac{1}{1+s^2}.$$

Hence

$$\tan^{-1}x + \tan^{-1}\left(\frac{1}{x}\right) = \int_0^x dt\, \frac{1}{1+t^2} + \int_x^{\infty} ds\, \frac{1}{1+s^2} = \int_0^{\infty} dt\, \frac{1}{1+t^2} = \frac{\pi}{2}.$$

Solution for Exercise 1.75 label: ex:vp1-65e

Differentiation gives $g'(x) = 2f(2x) - f(x)$ so $g'(x) = 0$ when $2f(2x) = f(x)$. In the first case this gives $2e^{2x} = e^x$ and hence $x = -\ln 2$ is the only real solution.

In the second case the equation becomes $\dfrac{\sin 2x}{x} = \dfrac{\sin x}{x}$. Since $x \neq 0$, the equation becomes $2\sin x\cos x = \sin x$, hence $\sin x = 0$, that is $x = n\pi$, $n = \pm1, \pm2, \cdots$, or $\cos x = 1/2$, that is $x = 2n\pi \pm \pi/3$.

Solution for Exercise 1.76 label: ex:vp1-66e

The definition 1.41 of an integral, with $x_k = a + kh/n$, $k = 1, 2, \cdots, n$ gives

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

(a) Put $f(x) = x^5$, $a = 0$ and $h = 1$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left(\frac{k}{n}\right)^5 = \int_0^1 dx\, x^5 \quad \text{hence} \quad \lim_{n\to\infty} \frac{1}{n^6} \sum_{k=1}^{n} k^5 = \frac{1}{6}.$$

(b) Put $f(x) = 1/(1+x)$, $a = 0$, $h = 1$ and sum from $k = 1$ to $k = 2n$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{2n} \frac{1}{1+k/n} = \int_0^2 dx\, \frac{1}{1+x} \quad \text{hence} \quad \lim_{n\to\infty} \sum_{k=1}^{2n} \frac{1}{n+k} = \int_0^2 dx\, \frac{1}{1+x} = \ln 3.$$

(c) Consider the complex sum, with $f = e^{ixy}$, $a = 0$ and $h = 1$,

$$S = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} e^{iky/n} = \int_0^1 dx\, e^{ixy}$$

Hence

$$S = \frac{1}{iy}\left(e^{iy} - 1\right) = \frac{\sin y}{y} + \frac{2i}{y} \sin^2\left(\frac{y}{2}\right)$$

and hence

$$\lim_{n\to\infty} \frac{1}{n} \left(\sin\left(\frac{y}{n}\right) + \sin\left(\frac{2y}{n}\right) + \cdots + \sin y \right) = \frac{2}{y} \sin^2\left(\frac{y}{2}\right).$$

(d) If $P_n = \frac{1}{n}\Big[(n+1)(n+2)\dots(2n)\Big]^{1/n}$ then

$$\ln P_n = -\ln n + \frac{1}{n} \sum_{k=1}^{n} \ln(k+n) = \frac{1}{n} \sum_{k=1}^{n} \ln(1+k/n).$$

But, with $f(x) = \ln(1+x)$, $a = 0$ and $h = 1$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \ln(1+k/n) = \int_0^1 dx\, \ln(1+x) = \ln 4 - 1.$$

Hence $\lim_{n\to\infty} P_n = \exp(\ln 4 - 1) = 4/e$.

label: ex:vp1-67e

Solution for Exercise 1.77

In these two cases the formula 1.47 does not work, because the integrand is infinite at $x = u$. However, it is clear that both $F'(u)$ and $G'(u)$ exist in some cases, for instance $f = g = 1$ or x, so the equivalent of expression 1.47 ought to exist.

In the first case the simplest method is to remove the singularity at $x = u$ using the standard change of variable $x = u\sin\phi$ to give

$$F(u) = \int_0^{\pi/2} d\phi\, f(u\sin\phi).$$

Now we can use equation 1.47 to give

$$F'(u) = \int_0^{\pi/2} d\phi \, \frac{\partial}{\partial u} f(u \sin\phi) = \int_0^{\pi/2} d\phi \, f'(u \sin\phi) \sin\phi.$$

This, second expression, may be converted back to an integral over x,

$$F'(u) = \frac{1}{u} \int_0^u dx \, \frac{x f'(u)}{\sqrt{u^2 - x^2}}.$$

In the second case we use another, more general trick. Consider the integral

$$G_\delta(u) = \int_0^{u-\delta} dx \, \frac{g(x)}{(u-x)^a}, \quad \delta \geq 0,$$

for which equation 1.47 is valid, when $\delta > 0$. This gives

$$G'_\delta(u) = \frac{f(u-\delta)}{\delta^a} + \int_0^{u-\delta} dx \, f(x) \frac{\partial}{\partial u} \left(\frac{1}{(u-x)^a} \right), \quad \delta > 0.$$

Now write $\frac{\partial}{\partial u} \left(\frac{1}{(u-x)^a} \right) = -\frac{\partial}{\partial x} \left(\frac{1}{(u-x)^a} \right)$ and integrate by parts

$$\begin{aligned} G'_\delta(u) &= \frac{f(u-\delta)}{\delta^a} - \left[\frac{f(x)}{(u-x)^a} \right]_0^{u-\delta} + \int_0^{u-\delta} dx \, \frac{f'(x)}{(u-x)^a}, \\ &= \frac{f(0)}{u^a} + \int_0^{u-\delta} dx \, \frac{f'(x)}{(u-x)^a}. \end{aligned}$$

Now take the limit $\delta \to 0$ to obtain

$$G'(u) = \frac{f(0)}{u^a} + \int_0^u dx \, \frac{f'(x)}{(u-x)^a}.$$