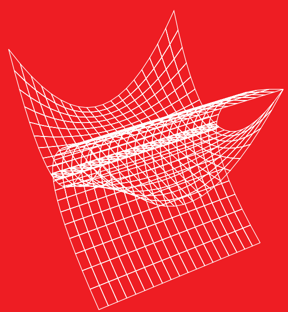
MathsWorks for Teachers *Series editor* David Leigh-Lancaster



Functional Equations

David Leigh-Lancaster

This book provides mathematics teachers with an introduction to elementary aspects of functional equations. These equations are linked to function in various topics of the senior secondary mathematics curriculum including transformations, identities, difference equations and mathematical modelling. A computer algebra system has been used to generate tables and graphs, as well as carrying out symbolic computation for the illustrative examples.

Functional equations are equations that have functions as solutions. They have been studied in some form or other since antiquity, and especially from the 19th century. At this time, further consideration was given to the notion of function in general, and in-depth analysis of functions, derivatives, integrals and their properties. The contemporary study of functional equations in mathematics involves their widespread use to model a broad range of practical and theoretical situations. Functional equations provide a powerful and concise means of characterising the algebraic properties of functions, in particular identities, and the use of mathematical software for computation.

Series overview

MathsWorks for Teachers has been developed to provide a coherent and contemporary framework for conceptualising and implementing aspects of middle and senior mathematics curricula.

Titles in the series are:

Functional Equations
David Leigh-Lancaster

Contemporary Calculus
Michael Evans

Matrices
Pam Norton

Foundation Numeracy in Context
Gary Motteram & David Tout

Data Analysis Applications
Kay Lipson

Complex Numbers and Vectors
Les Evans

ISBN 0-86431-492-2
9 780864 314925

Australian Council *for* Educational Research
ACER

MathsWorks for Teachers
FUNCTIONAL EQUATIONS David Leigh-Lancaster

Functional Equations

David Leigh-Lancaster

MathsWorks for Teachers

Functional Equations

David Leigh-Lancaster

MathsWorks for Teachers

First published 2006
by ACER Press
Australian Council for Educational Research Ltd
19 Prospect Hill Road, Camberwell, Victoria, 3124

Edited by Ruth Siems Editorial Services
Cover design by FOUNDRY Typography, Design & Visual Dialogue
Text design by Robert Klinkhamer
Typeset by Desktop Concepts P/L, Melbourne
Printed by Shannon Books Australia Pty Ltd

National Library of Australia Cataloguing-in-Publication data:

Leigh-Lancaster, David.
Functional equations.

Bibliography.
ISBN 0 86431 492 2.

1. Functional equations. I. Title.

515.75

Visit our website: www.acerpress.com.au

CONTENTS

INTRODUCTION

MathsWorks is a series of *teacher* texts covering various areas of study and topics relevant to senior secondary mathematics courses. The series has been specifically developed for teachers to cover helpful mathematical background, and is written in an informal discussion style.

The series consists of six titles:

- An Introduction to Functional Equations
- Contemporary Calculus
- Matrices
- Data Analysis Applications
- Foundation Numeracy in Context
- Complex Numbers and Vectors

Each text includes historical and background material; discussion of key concepts, skills and processes; commentary on teaching and learning approaches; comprehensive illustrative examples with related tables, graphs and diagrams throughout; references for each chapter (text and web-based); student activities and sample solution notes; and a bibliography.

The use of technology is incorporated as applicable in each text, and a general curriculum link between chapters of each text and Australian state and territory as well as and selected overseas courses is provided.

A Notes section has been provided at the end of the text for teachers to include their own comments, annotations and observations. It could also be used to record additional resources, references and websites.

ABOUT THE AUTHOR

David Leigh-Lancaster is an experienced mathematics educator who has been a head of faculty and teacher of senior secondary mathematics for many years. His mathematical interests and background are in the areas of mathematical logic, computability theory and the history, foundations, and philosophy of mathematics. He has worked extensively in curriculum development and assessment, resource and teacher professional development and mathematics education research. He has a particular interest in the use of technology in mathematics teaching and learning.

CHAPTER 1

EQUATIONS, FUNCTIONS AND ALGEBRA

There are lots of different kinds of equations, so what are *functional* equations? Senior secondary mathematics students and their teachers would be familiar with various situations that involve equations and functions, such as solving an equation of the form $f(x) = 0$ to find the horizontal axis intercepts of the corresponding graph, or solving an equation of the form $f(x) = g(x)$ to find the coordinates of the points of intersection of the graphs of the two functions, and then possibly to determine the area between the two curves.

For example, if $f: R \rightarrow R$, $f(x) = x^4$ and $g: R \rightarrow R$, $g(x) = x^2$, then the graphs of the two functions, as shown in Figure 1.1, intersect when $f(x) = g(x)$, that is when $x^4 = x^2$, or alternatively, when $f(x) - g(x) = 0$, that is $x^4 - x^2 = 0$. The roots of this equation are readily determined as $\{x: x = -1, x = 0, x = 1\}$. The solutions to these sorts of equations are sets of real numbers, possibly even the *empty set* when the equation has *no solution*.

```
f[x_] : = x^4     g[x_] : = x^2
Plot[{f[x], g[x]}, {x, -3, 3}, AxesLabel → {x, y}, PlotRange → {-1, 6}]
```

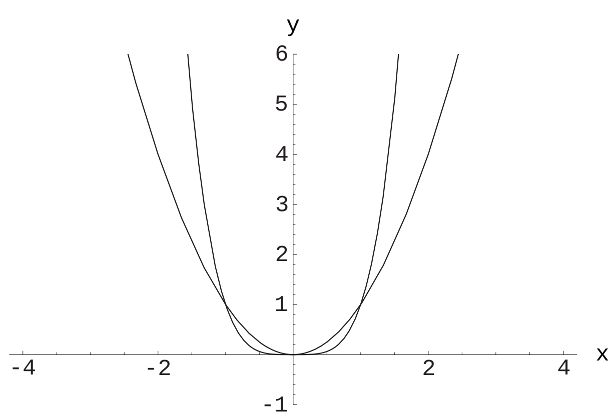


Figure 1.1: Graphs of *f* and *g*

The area between the graphs of the two functions and their points of intersection, as shown in Figure 1.2, can be found by evaluating a suitable definite integral. From the graphs in Figure 1.1 it can be seen that both graphs are symmetrical about the y-axis and that $g(x) \geq f(x)$ on the interval [–1, 1].

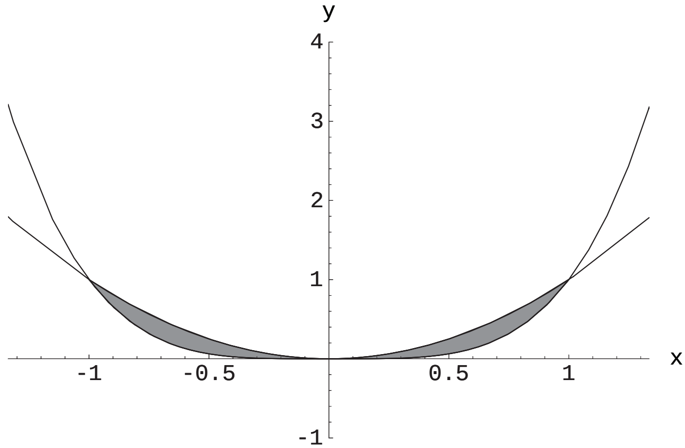


Figure 1.2: Graphs of *f* and *g* and area between two curves and their points of intersection

Thus, the required area can be obtained by evaluating the definite integral:

$$2\int_0^1 (g[x] - f[x])\, dx$$

$$\frac{4}{15}$$

Both of the functions f and g are many-to-one functions, and therefore do not have inverse functions. If their domains are restricted to the interval $[0, \infty)$, they become one-to-one functions. To find the functions f^{-1} and g^{-1} such that:

$$f^{-1}(f(x)) = x = f(f^{-1}(x)) \quad \text{and} \quad g^{-1}(g(x)) = x = g(g^{-1}(x))$$

respectively, is to seek in each case a *function* which is a solution of an equation. These two equations are thus simple *functional* equations. They also happen to characterise a particular *symmetry* property, reflection symmetry in the line $y = x$, as shown for the graph of g in Figure 1.3.

```
Plot [{g[x], x, √x}, {x, 0, 9}, PlotRange → {-1, 10}
   AspectRatio → Automatic, AxesLabel → {x, y}]
```

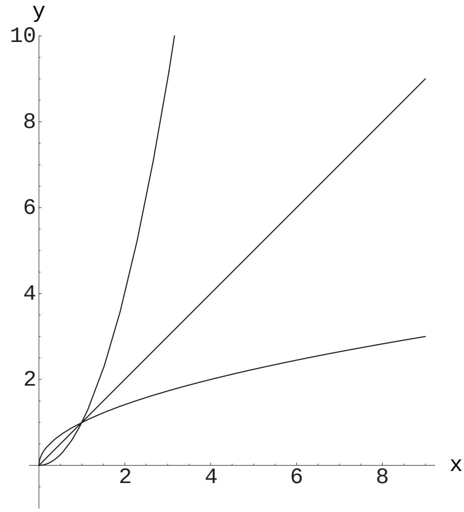


Figure 1.3: Graphs of restricted function g, the line $y = x$ and g^{-1}

Functional equations are *equations*, whose variables and *solutions* are *functions* rather than numbers. Like most straightforward descriptions, this raises several other related questions:

- Why are *functional* equations of interest?
- What are the sorts of mathematical *problems* and *techniques* associated with functional equations?
- How can the study of functional equations be *related* to the senior secondary mathematics *curriculum*?

This resource is designed to provide an introductory response, at least in part, to each of these questions. The contemporary study of functional equations and their applications to mathematical modelling in the life sciences, natural sciences, economics and business proceeds from the mid 1960s, following the earlier development of a more comprehensive understanding of the concept and theory of functions, and computation and function in the 19th century and early 20th century.

Functional equations underpin the study of several key aspects of the mathematics curriculum, yet their role in this is rarely made explicit, partly because it is often subsumed within various well-known 'topics' and partly because it requires consideration of functions themselves as 'objects' for investigation. That is, for teachers and students to consider *function* as a *reified construct* in its own right (from the Latin *re* for *thing*). To facilitate such mathematical inquiry, and consideration of its application to mathematics curriculum, it is helpful to look at some key background ideas such as *equality*, *equation*, *function* and *graphs*, *algebra* and *computation*. That is the purpose of this chapter.

The study of functions has been much aided by the development of mathematically able mechanical technologies in the 19th century (for example, the work of Charles Babbage) and up to the 1930s, and the subsequent development of first electro-mechanical (1937) and then electronic technologies from 1947 to the modern digital computer.

These now support readily available software such as spreadsheets, graphics calculators and computer algebra systems (CAS) that enable functions to be defined as objects, which are then manipulated, used in equations and systems of equations, transformed, and the like. CAS in particular support natural links between numerical, graphical and symbolic representation and computation.

For example, if the function h is defined in terms of f and g by $h: R \to R$, $h(x) = f(x) - g(x) = x^4 - x^2$, then its rate of change or derivative function can be represented graphically as shown in Figure 1.4.

```
Plot [h' [x], {x, - 2}, AxesLabel → {x, "h'[x]"}] :
```

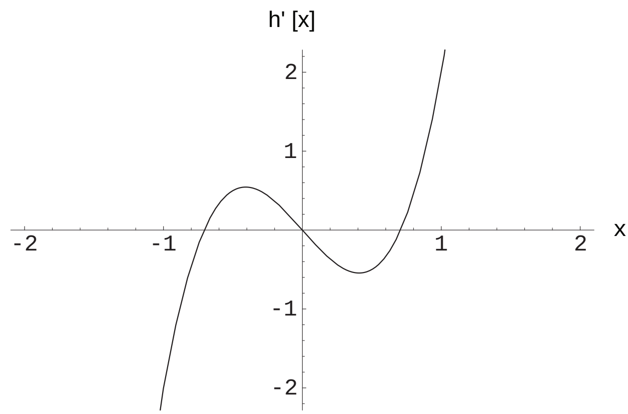


Figure 1.4: Graph of *h′* the derivative function of *h*

Using CAS, the derivative at a point can be evaluated *numerically*, for example at $x = 1$:

$$N\left[\frac{h[1 + 10^{-6}] - h[1]}{10^{-6}}, 20\right]$$

2.0000050000040000010

or it can be evaluated from *first principles* if so desired:

$$\text{Limit}\left[\frac{h[1 + \delta x] - h[1]}{\delta x}, \delta x \to 0\right]$$

2

or determined *symbolically* using the application of the CAS routine for symbolic differentiation:

```
h'[1]
2
```

There are different hand-held and computer-based platforms for such technologies, and CAS are indispensable tools for the thorough investigation of functions and related topics in mathematics. Although the material in this resource has been developed using the CAS *Mathematica*, such use is not intended as an instructional course in the use of this particular CAS. Rather, it is intended to illustrate various aspects of working mathematically with functions, their graphs and related equations with the assistance of such technology.

Important CAS functionality such as **`Plot`**, **`Solve`**, **`Differentiate`** and the like is typically generic in form across different implementations, and in each case designed to closely model standard mathematical forms and conventions. Thus, while there are minor variations between different CAS, teachers and students will be able to fairly readily identify the underpinning computational constructs and related functionality. Indeed, there is increasing convergence between hand-held, palmtop–laptop and desktop platforms for all kinds of technology and software functionality. Recent developments in graphics calculators technology (2000–2005) with memory capacities of several megabytes and an increasingly sophisticated range of quasi-CAS supplementary programs also see convergence in functionality from this technology as well.

EQUALS AND EQUALITY

Students meet, and become familiar with, the notion of *equality* from early on in their study of mathematics, although they may not initially be explicitly aware of this. Along with numbers and arithmetic operations, the '=' symbol and the corresponding word 'equals' is one of the early mathematical symbols that students learn to recognise and apply, although it was not introduced into mathematics until 1557 through the work of Roberte Recorde in his *Whetstone of Witte.* Alternative representation such as *ae*, (from the Latin *aequalis* or *equal*) and || were still in common use in the 1700s.

From the early years of schooling, students move from informal conceptions and uses of the notion of 'equals' through to formal representations and related numerical, logical and algebraic techniques as they progress through the compulsory years, and subsequently senior secondary years, of school mathematics education.

The expression *equals* enters mathematical language at an early stage in schooling, where the notion of *equality* is usually expressed initially in terms of the verb *to be*. For example, the sense of 'number' is related to 'equal groups' of objects, connected by the notion of *a one-to-one correspondence* between sets of the 'same size'—which is itself a *function* (see Skemp 1989, for a detailed discussion of the early stages of student mathematics learning).

Early student arithmetic statements are usually expressed verbally, or in written form in natural language, using forms such as 'one and one *is* two' or 'three fours *are* twelve'. These are convenient and commonly used abbreviations for 'one and one *is equal to* two' and 'three fours *are equal to* twelve' respectively, or, in symbolic form '$1 + 1 = 2$' and '$3 \times 4 = 12$'. However, this informality can disguise the fact that the mathematical expression is quite precise in its meaning compared with the verb 'to be' which can be used in several ways. For example, the statement 'Tessa *is* a girl' is not intended to identify the object 'girl' with another object 'Tessa' but rather to indicate that *Tessa* is a member (or element) of the set of *girls*. Here the role of the verb 'is' is as an *existential quantifier*—it asserts the existence of an element (Tessa) of a set (girls). Mathematically, this can be expressed as $Tessa \in G$, where G is the set of all girls, and the symbol $\in$ (is an element of) designates the set relation of membership.

The notion of *number* is both subtle and abstract, and, from one perspective, may be thought of as denoting the *size* of a set. The *symbol* '**12**' is a numerical designation, or *numeral*, for a certain number in the Hindu-Arithmetic numeration system, just as the symbol '**XII**' is the corresponding designation in the Roman numeration system. While the numerical designation of a number can change from one language to another, the *mathematical* properties of the number do not. Thus, regardless of the language and particular symbols used in mathematics the result of the standard computation: 'three multiplied by four' is *always* equal to 'twelve'.

On the other hand, there are infinitely many possible computations that yield 'twelve' as a result. For example, while '$3 \times 4 = 12$' it is also the case that '$2 \times 6 = 12$'; '$6 + 6 = 12$'; '$3 + 9 = 12$' and '$14 - 2 = 12$' … and so on. The part of the expression on the left of the '=' symbol is different in each case, but the part of the expression on the right of the '=' symbol is the same in each case. Thus, the *result* of each of these computations is the *same number*. In mathematics '=' is a special type of relation called an *equivalence* relation. The relation of equality may have different interpretations, for example, between sets its is usually taken to mean that two sets are equal if, and only if, they have exactly the same elements; while in geometry equality can be interpreted in terms of *congruency* of shapes.

To say that equality is an equivalence relation means that if x, y and z are arbitrary elements of some set S on which this relation is defined (such as the set of natural numbers, $N = \{0, 1, 2, \ldots\}$, for which the relation '=' has its usual interpretation) then:

- $x = x$ reflexivity
- $x = y$ implies $y = x$ symmetry
- $x = y$ and $y = z$ implies $x = z$ transitivity

The symbols x, y and z that represent *arbitrary* elements of the set S are called *variables*. They may take any value in S, and although they are usually distinct elements they do not need to be, thus cases where some or all of x, y, z are the same element are automatically included.

These properties may seem fairly evident and this is the case where it is *true* that a particular relation happen to be an *equivalence relation* on a given set. However, this is usually the point in question, as there are many different relations that can be defined on a given set, some of which happen to be equivalence relations, while others are not.

For example, a familiar relation which is *not* an equivalence relation on the set of natural numbers is the relation 'less than' or '<'. This simplest way to show this is to give a counter-example to at least one of the defining conditions for an equivalence relation:

- $x < x$ is **false** for $x \in N$: for example, 5 is not less than 5
- $x < y$ does **not** imply $y < x$ for $x, y \in N$: for example, $3 < 4$ does not imply $4 < 3$

however:

- $x < y$ and $y < z$ implies $x < z$ is **true** for $x, y, z \in N$

Thus '<' is *not* an equivalence relation, since the first two conditions do *not* hold, although the third condition does.

This illustrates an important principle in mathematical reasoning—it is sufficient to exhibit a *single* counter-example to a statement to show that it is *not* true, or, in other words, is *false*. To show that something is true in general usually requires some form of reasoning using a free (arbitrary) variable, such as is commonly used in algebraic proofs. In work with functions this typically involves some form of analytical reasoning.

For natural numbers, N, the relation $x < y$ can be defined in terms of the relation '=' by the logical statement that asserts the *existence* of a unique *non-zero* natural k such that $x + k = y$. Thus, $6 < 10$ since there exists a non-zero natural number, **4**, where 6 + **4** = 10. However, it is *not* the case that 6 < 3, since there is *no* natural number k such that $6 + k = 3$.

Students also progressively become familiar with other properties of '=' during their study of mathematics, in particular where computation with number is involved directly, but also with respect to certain so called 'algebraic properties' or 'principles of algebraic manipulation'. For example, if x and y are natural numbers such that $x = y$, then for any natural number z, the following are also true:

- $x + z = y + z$
- $x \times z = y \times z$

Again, these may seem evident; however, they include specific cases such as:

- $(3 \times 4) + 5 = (2 \times 6) + 5$
- $(3 + 9) \times 4 = (14 - 2) \times 4$

as well many other substitution instances where the initial equality holds. This may not be immediately apparent unless such examples are considered explicitly. A detailed and accessible discussion on the structure of number can be found in Stillwell (1999).

In senior secondary mathematics, students who undertake courses based on the study of function, algebra, coordinate geometry and calculus often begin these by some study of polynomial functions of a single real variable. This is typically developed as a natural extension of earlier study involving linear and quadratic polynomial functions. Thus, a general polynomial function, p, of degree n, is defined as

$$p: R \to R\ ,\ p(x) = a_n x^n + a_{n-1}\, x^{\,n-1} + a_{n-2}\, x^{n-2} + \ldots + a_2\, x^2 + a_1\, x^1 + a_0 x^0.$$

If q is another polynomial function, also of degree n,

$$q: R \to R,\ q(x) = b_n x^n + b_{n-1}\, x^{n-1} + b_{n-2}\, x^{n-2} + \ldots + b_2\, x^2 + b_1\, x^1 + b_0 x^0,$$

then $p = q$ if, and only if, all the corresponding coefficients are equal, that is:

$$a_n = b_n,\ a_{n-1} = b_{n-1},\ \ldots,\ a_2 = b_2,\ a_1 = b_1 \text{ and } a_0 = b_0.$$

Many students regard this as an 'obvious' statement, but do not necessarily connect it to the process for re-expression of a polynomial expression $p(x)$ with respect to a given linear term $(x - a)$ as $p(x) = (x - a)\, q(x) + r$, where in the case of a cubic polynomial function, $q(x)$ is a quadratic polynomial function and r is a real number. Yet understanding of the equivalence of these forms through this re-expression underpins the remainder and factor theorems for polynomial functions. In most cases that students deal with, the coefficients a_i

and b_i are integers, and the re-expression process can be carried out more or less readily by application of the long division algorithm from arithmetic, or directly by equating coefficients.

For example, if $p(x) = x^3 + 2x^2 - 3x + 1$ and the linear term is $(x - 2)$ then, equating coefficients from one term to the next, and ensuring that the coefficients match up with those of the original expression gives:

$$\begin{aligned} p(x) &= x^2 (x-2) + \ldots(x-2) + \ldots(x-2) + \ldots && \text{(to get the } x^3 \text{ term)} \\ &= x^2 (x-2) + 4x(x-2) + \ldots(x-2) + \ldots && \text{(to balance } -2x^2 \text{ and get } +2x^2\text{)} \\ &= x^2 (x-2) + 4x(x-2) + 5(x-2) + \ldots && \text{(to balance } -8x \text{ and get } -3x\text{)} \\ &= x^2 (x-2) + 4x(x-2) + 5(x-2) + 11 && \text{(to balance } -10 \text{ and get } +1\text{)} \\ &= (x-2)(x^2 + 4x + 5) + 11 && \text{(collecting } x-2 \text{ terms)} \end{aligned}$$

Similarly, if the linear term is $(x + 1)$, the same process can be carried out again, written in this instance more concisely in two lines of working , one for the re-expression and one for collecting $x + 1$ terms.
This gives
$p(x) = (x - a)\ q(x) + r$ as:

$$\begin{aligned} p(x) &= x^2 (x+1) + x(x+1) - 4(x+1) + 5 \\ &= (x+1)(x^2 + x - 4) + 5 \end{aligned}$$

Many students do not realise that the 'choice' of the linear term is completely *arbitrary*, the procedure is *algorithmic* and can be carried out just as well for $(x + 3\ )$ or $(x - 4)$, or any other selection for this term:

$$\begin{aligned} p(x) &= x^2 (x+3) - x(x+3) + 0(x+3) + 1 \\ &= (x+3)(x^2 - x) + 1 \end{aligned}$$

$$\begin{aligned} p(x) &= x^2 (x-4) + 6x(x-4) + 21(x-4) + 85 \\ &= (x-4)(x^2 + 6x + 21) + 85 \end{aligned}$$

The equality of polynomial functions is an equivalence relationship, so any one of these forms $p(x) = (x - a)\ q(x) + r$ for the rule of the cubic polynomial function p is equal to any other of the forms $p(x) = (x - a)\ q(x) + r$ irrespective of the value of a.

A consequence of this is that for any particular value of the variable x, the value of $p(x)$ will be the same. In each re-expression form of this rule it can readily be verified that $p(0) = 1$ and that $p(2) = 11$. Table 1.1 shows the equivalence the first three forms for integer values of x from –5 to 5.

Table 1.1: Equivalence of forms of *p*(*x*) for integer values of *x* from –5 to 5

```
p1[x_]: = x^3 + 2x^2 - 3x + 1
p2[x_]: = (x - 2)(x^2 + 4x + 5) + 11
p3[x_]: = (x + 1)(x^2 + x - 4) + 5
Table[{x, p1[x_], p2[x_], p3[x_]}, {x, -5, 5}]
```

-5	-59	-59	-59
-4	-19	-19	-19
-3	1	1	1
-2	7	7	7
-1	5	5	5
0	1	1	1
2	11	11	11
3	37	37	37
4	85	85	85
5	161	161	161

However, the general reasoning for the *equivalence* of these expressions over the natural domain of the function, R, requires algebraic manipulation within the relevant mathematical structure, the ring of polynomials.

Functions are a special type of *set* that can be represented using ordered pairs, rules and graphs. Two *sets* are *equal* if and only if they are comprised of exactly the same elements. In general, two functions, f and g, are said to be equal if and only if:

- they have the same domain, that is, $d_f = d_g$; and
- $f(x) = g(x)$ for all x in $d_f = d_g$.

Another way of saying this is that $\{(x, f(x)), x \in d_f\} = \{(x, g(x)), x \in d_g\}$, that is, the *sets of ordered pairs* that represent the two functions are equal. Hence the *graphs* of the two functions will also be identical (coincide) if the two functions are equal. Thus the functions

$$f: R \to R, f(x) = x^2 - 2x - 5 \quad \text{and} \quad g: R \to R, g(x) = (x-1)^2 - 6$$

are equal and have the same graph over a given interval, as illustrated in Figures 1.5a and 1.5b:

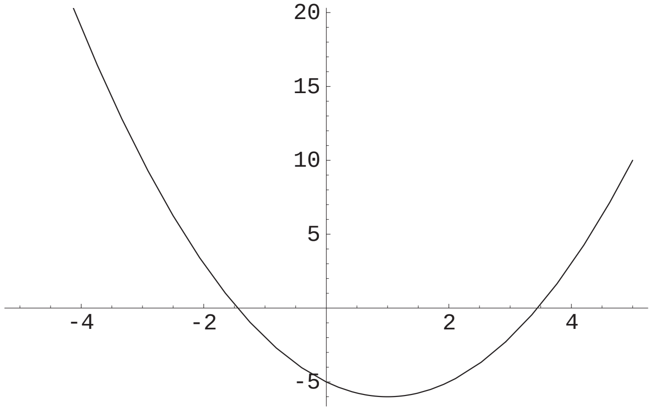


Figure 1.5a: Graph of f: $R \to R$, $f(x) = x^2 - 2x - 5$

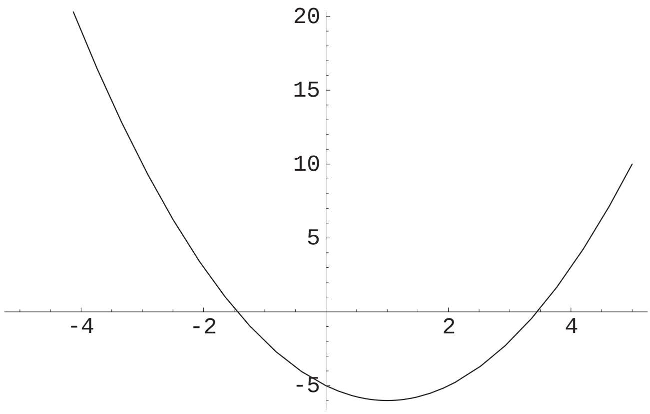


Figure 1.5b: Graph of g: $R \to R$, $g(x) = (x - 1)^2 - 6$

That is, in this case, the set of ordered pairs $\{(x, y): x \in R \text{ and } y = f(x)\}$ is the same as (is equal to) the set of ordered pairs $\{(x, y): x \in R \text{ and } y = g(x)\}$.

However, if f remains the same and the function g is redefined so that:

$$f: R \to R,\ f(x) = x^2 - 2x - 5 \quad \text{and} \quad g: Z \to R,\ g(x) = x^2 - 2x - 5$$

then $f \neq g$, since while they have the same rule they do *not* have the *same domain*. The graph of f is a continuous curve and the graph of g is a set of discrete points, as shown in Figure 1.6:

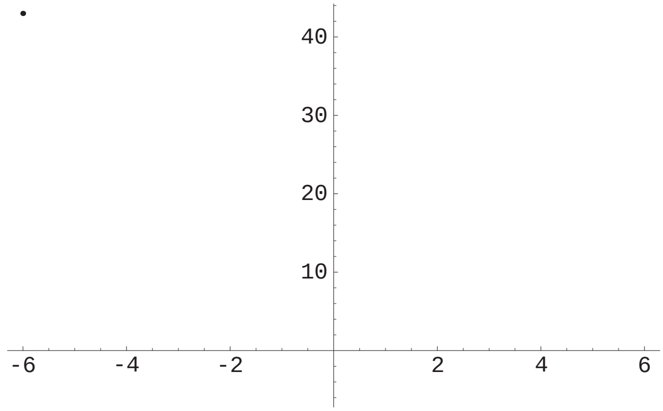


Figure 1.6: Graph of g: $Z \to R$, $g(x) = x^2 - 2x - 5 = (x - 1)^2 - 6$

In this case, the set of ordered pairs that specifies the function g is an infinite and discrete *proper subset* of the set of ordered pairs that specifies the function f. That is:

$$\{(x, y): x \in Z \text{ and } y = g(x)\} \subset \{(x, y): x \in R \text{ and } y = f(x)\}$$

Similarly, if f remains the same and the function g is redefined so that:

$$f: R \to R, f(x) = x^2 - 2x - 5 \quad \text{and} \quad g: R \to R, g(x) = (x - 1)^2 + 4$$

then $f \neq g$, since while they have the same domain they do *not* have the same *range*. For example, $(1, -6) \in f$ whereas $(1, 4) \in g$. The graphs of the these two functions do not coincide, as shown in Figure 1.7:

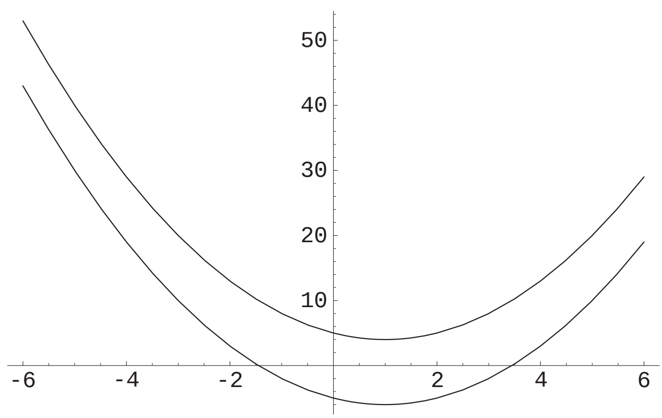


Figure 1.7: Graph of f: $R \to R$, $f(x) = x^2 - 2x - 5$ and g: $R \to R$, $g(x) = (x - 1)^2 + 4$

If two functions f and g are specified by domain and rule, and have the *same domain*, then they will be equal if the *difference* between $f(x)$ and $g(x)$ is always zero on their common domain, that is $f(x) - g(x) = 0$. CAS may be of assistance in determining whether this is the case or not, especially where the form of the rules for f and/or g is complicated. In simpler cases this can be determined by hand. For example, if :

$$f: R \rightarrow R, f(x) = 4x - 7 \quad \text{and} \quad g: R \rightarrow R, g(x) = 2(2x - 3) + 1,$$

then

$$f(x) - g(x) = 4x - 7 - (2(2x - 3) + 1) = 4x - 7 - 4x + 6 - 1 = -2 \neq 0, \text{ and so } f \neq g.$$

However, if:

$$g: R \rightarrow R, f(x) = 2(2x - 4) + 1$$

then

$$f(x) - g(x) = 4x - 7 - (2(2x - 4) + 1) = 4x - 7 - 4x + 8 - 1 = 0, \text{ and so } f = g.$$

EQUATIONS

A simple description of an 'equation' might be any 'well-defined' mathematical expression that contains the symbol '=', or a natural language equivalent in words. Early work on equations by students involves what are sometimes called quasi-variables (see Stephens 2004). These are informal versions of variables, and may not be denoted by a symbol or other term, but can be kept in mind as signifying a sense of arbitrary designation.

Students are often asked to identify numbers that make certain mathematical (arithmetic) statements true. At first this typically involves situations where only one value exists such as 'fill the space '…' with the missing number that makes the number sentence '4 + … = 12' true, and often occurs in contexts where the operation of subtraction is being considered. Indeed, later on in a student's mathematical studies, the insufficiency of the natural numbers to provide truth values for similar looking statements such as ' 9 + … = 5' is used as a motivation for the introduction of the set of integers, $Z = \{ \dots -3, -2, -1, 0, 1, 2, 3 \dots \}$. Sometimes 'small print' geometric shapes and symbols such as ♣ or 'empty' shapes such as triangles and squares are used to stand for the 'missing number', although this can introduce ambiguities of its own.

For example, does '□' in the expression ' 9 + □ = 5' stand for a box that is intended to be 'filled in' with the number −4, such as:

$$9 + \boxed{-4} = 5$$

or is it intended as a symbol for a quasi-variable where □ is *itself* taken to be –4, that is, □ = –4?

Where such forms are used, care will need to be taken to ensure that students understand in what sense the use of a given symbol is intended, and in which context. Perhaps it is for this reason that literal symbols offer some advantage, although they have their own complications, being variously used to represent: *constants*; *undetermined coefficients* (but which are actually implicitly understood to have specific values); and *variables*, or the *value(s)* of a variable or variables that satisfy the conditions of equations and/or other conditions.

Thus, students should be aware that while an equation over the domain of natural numbers, N, such as $a + 7 = 10$ has a *unique* solution $a = 3$, the equation $a + b = 10$ has a *finite* set of multiple solutions over the same domain, including $a = b = 5$ as well as $a = 8$ and $b = 2$; and has an *infinite* set of solutions of the form $a, b = 10 - a$, including, for example, $a = 16$ and $b = 10 - 16 = -6$ over the domain of integers, Z.

As students progress to include consideration of many-to-one functions in their studies, such as quadratic functions, other polynomial functions of higher degree and circular functions, they further develop their understanding of the idea that, for some functions, an equation of the form $f(x) = a$ has multiple truth values over the domain of the function, or a given subset of this domain.

Indeed, in the case of circular functions, there may be no solutions, for example, $2\sin(3x) = 5$, as shown in Figure 1.8,

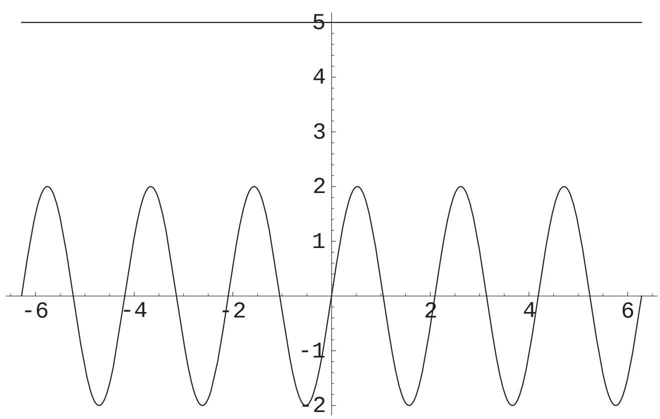


Figure 1.8: Part of the graphs of $f(x) = 2\sin(3x)$ and $y = 5$

or there may be finitely many solutions, as shown in Figure 1.9 for the equation $2\sin(3x) = 1$ over the interval $[-2\pi, 2\pi]$:

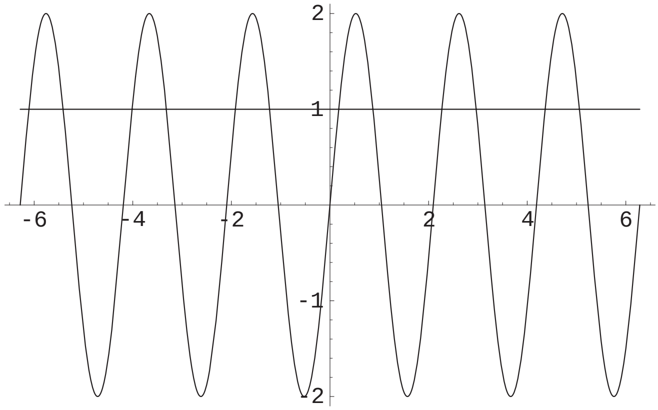


Figure 1.9: Graphs of $f(x) = 2\sin(3x)$ and $y = 1$ over the interval $[-2\pi, 2\pi]$

In other cases there may be infinitely many solutions over the natural domain of the function. This can only be *indicated* graphically, by consideration of the infinite extensibility of the graphs of $y = 1$ and $f(x) = 2\sin(3x)$ over R, and the periodic nature of the graph of $f(x) = 2\sin(3x)$ as observed for several cycles, as illustrated in Figure 1.10:

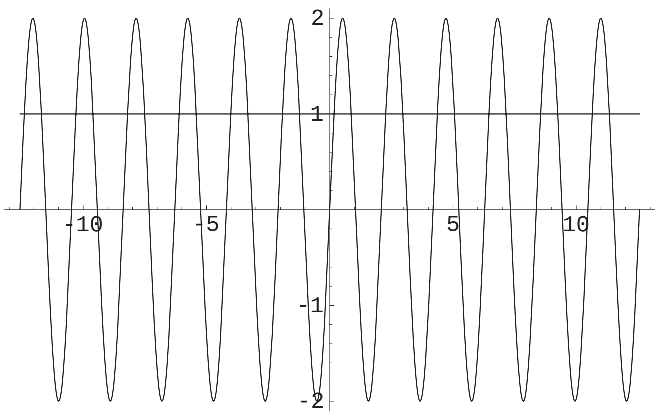


Figure 1.10: Graphs of $f(x) = 2\sin(3x)$ and $y = 1$ for several periods of f

Sometimes equations involving functions can be solved analytically using algebraic techniques; in other cases graphical or numerical techniques need to be employed to find approximate solutions to a required accuracy, over a given interval. When a many-to-one function is involved, a *principal domain* solution for a one-to-one subset of the natural domain is often used, and multiple solutions over the larger domain generated from this value.

FUNCTION, ALGEBRA AND COMPUTATION

Functions are part of the lifeblood of mathematics, link to both computation and algebra, and involve the notion of equality in both contexts. In work on coordinate geometry and calculus, functions are often described in terms of some sort of rule that assigns (maps) elements of one set, X, to elements of another set, Y, with the defining property that each element in the set X is assigned (mapped) to exactly one element in the set Y.

If x and y are selected as variables to represent elements of X and Y respectively, then for a function, f, from X to Y, the mapping can be denoted by: $x \to y$, $x \to f(x)$ or $y = f(x)$. In this case x is called the *independent* variable and y is called the *dependent* variable. A function can be defined simply as a set of ordered pairs in which no ordered pair in the set has the same first element as any other ordered pair in the set. If a rule of the form $y = f(x)$ is available to describe the mapping, then the function f can be defined by $f = \{(x, y): y = f(x) \text{ and } x \in X\}$.

What is commonly called 'the graph' of a function is its (usually partial) representation by points on a cartesian coordinate system (two axes at perpendicular to each other through a fixed reference point called the *origin*) where each ordered pair corresponds to the coordinates of a point. The axes do not need to be perpendicular, although most work on graphs in schools follows on from the work of Descartes in the 17th century and mainly uses the *cartesian* coordinate system. For example, two parallel lines could be used to form what is called an *arrow* (*mapping*) *diagram*, as illustrated in Figure 1.11 for the function $g: R \to R$, $g(x) = x^2$:

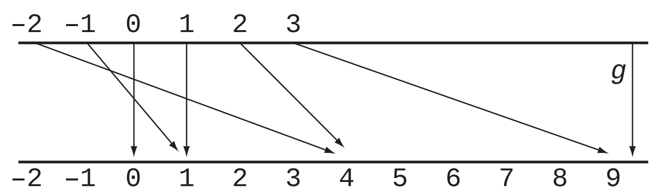


Figure 1.11: Part of an arrow (mapping) diagram for the function $g: R \to R$, $g(x) = x^2$

In general, for a function, f, from X to Y the set X is called the *domain* of f, written dom(f) or d_f, and the set Y is called the *co-domain* of f. Actually, many 'graphs' of functions, such as those shown earlier, are only plots of a *subset* of the function, but it is conventional practice to accept that, with experience, these subsets are in most cases sufficient to provide an adequate visual illustration of the 'behaviour' of the function. It may or may not be the case that every y in Y has an x from X assigned to it; the set of $f(x)$ values resulting from the application of $f(x)$ to all x in X is called the *range* of f, written ran(f) or r_f, and is a subset of Y, that is, $r_f \subseteq Y$.

In mathematical work involving functions and algebra, the mathematical *relation* of *equality* is used with several different meanings in different contexts, which frequently leads (in different ways) to challenges for students and teachers alike. The formal definition of '=' as an *equivalence relation* with certain properties that require *interpretation* with respect to the *mathematical structure* within which it is being used is part of the theory of mathematical logic (see Crossley 1972). This is a field of mathematics that evolved rapidly in the first half of the 20th century following problems and controversy in set theory and the foundations of mathematics (see Chaitin 2000). A closely related field of mathematics is that of computability theory, which draws strongly on mathematical logic and can be applied in any domain of mathematics where algorithms or heuristics are used, or sought, to solve particular classes of problems. These may include, for example, calculation of values of functions, the solution of equations, geometric constructions, or carrying out various computations. *Practical* computation *using technology* has required mathematicians to distinguish carefully between multiple uses of the symbol '=' for different purposes.

For example, the mathematical statement $f: R \to R$, $f(x) = ax + b$, where a and b are real constants, defines a family or class of functions of a given type, that is, linear functions of a single real variable. The particular function that is a member of this family, with rule $f(x) = 2x + 3$, is specified by assigning the value 2 to a and the value 3 to b, typically written as $a = 2$ and $b = 3$.

There are two distinct uses of '=' involved here. The first use of '=' is to define a relationship that holds between *variables* over a given set, the *domain* of the function, that is over *arbitrary* values of the domain. The second use of '=' is to *assign a fixed value* to the parameter a (defining it as a *constant*). If an equation such as $f(x) = 10$ is given, and solved for the variable x, then the particular values(s) of the variable that make the statement *true* are sought—which is a different use of the symbol '=' again.

Mathematically able software, such as computer algebra systems (CAS), distinguish between these uses of '=', sometimes explicitly using different symbols, sometimes by using the context of variable, formula or constant entry, or other devices.

Thus, while an expression such as $a = 2$, or $a \to 2$, is commonly used to assign a value to a constant, the symbol := is often used in programming languages to define the rule of a function, for example $f(x) := 2x + 3$. Some CAS such as *Mathematica*, which are used as programming languages as well as computational tools, distinguish explicitly between all three senses of '=', while others use devices such as special memory register designations. The following examples show how the CAS *Mathematica* assigns values to a *constant*, which can then be used in subsequent computations, how equations can be formulated and *solved*, and how a function can be *defined* and used to produce, tables and graphs.

CONSTANTS

Consider assigning the value 2 to the symbol a, that is $a = 2$, then various computations involving a will be performed with a acting as the constant value 2. For example, in the following **`bold courier new font`** expressions indicate *Mathematica* input, while `normal courier new font` expressions indicate *Mathematica* output:

```
a = 2
2
a + a + a
6
a^2 + 1
5
```

The assigned value of a will be retained unless it is explicitly cleared, then a returns to being an undetermined (free) symbol.

```
Clear[a]
a
a
a × a × a × a × a
a^5
```

Note that if the assigned value of a had been retained, the last computation would have the output 32.

This is a particularly important *practical consideration* for the use of graphics calculator or CAS technology. Literal symbols such as a, b, c are frequently used to represent coefficients, constants and the like; symbols such as f, g, h are used to represent functions, and symbols such as x, y, z to represent variables. At certain stages of working mathematically in a given context, it may well be that some or all of these are assigned specific values or definitions—in general, they will retain these unless they are cleared or redefined. Thus, if *a* is assigned the value 7, it cannot be used as an undetermined coefficient in later working *unless* its previous value has been cleared. It is important that students are asked to work through a selection of judiciously chosen examples to highlight the difference between a *variable* and a *constant*, which may or may not be assigned a specific value at a given time.

ALGEBRA AND EQUATIONS

When the solution of a particular equation is required, for example, finding the real values of x for which $2x^2 - 3x - 7 = 0$, it is *not* intended, as is the case with a constant, to assert that the algebraic expression $2x^2 - 3x - 7$ is to be regarded as a 'shorthand' for 0, but rather to find the particular values of x for which $2x^2 - 3x - 7$, when evaluated, will give the result 0. For the CAS *Mathematica*, a special symbol ' = =' is used to make this distinction:

Solve[2x^2 − 3x − 7 == 0, x]

$$\{\{x \to \frac{1}{4}(3 - \sqrt{65})\}, \{x \to \frac{1}{4}(3 + \sqrt{65})\}\}$$

Here the solution set is specified in terms of replacement values, which are interpreted as follows: if x is replaced by $\frac{1}{4}(3 - \sqrt{65})$ or if x is replaced by $\frac{1}{4}(3 + \sqrt{65})$ then evaluation of $2x^2 - 3x - 7$ will result in 0. The corresponding general solution for a quadratic equation with real coefficients and specification of any restrictions on defining parameters can be obtained by:

Reduce[ax^2 + bx + c == 0, x]

$$x == \frac{-b - \sqrt{b^2 - 4ac}}{2a} \;\&\&\; a \neq 0 \;||\; x == \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

$$a == 0 \;\&\&\; b == 0 \;\&\&\; c == 0 \;||\; a = 0 \;\&\&\; x == -\frac{c}{b} \;\&\&\; b \neq 0$$

where **&&** designates the logical connective 'and' and || designates the logical connective 'or'. Thus, the *Mathematica* functionality for **Reduce** provides all possible combinations of values for constants and variables that would lead to the stated equation being satisfied.

For the CAS *Mathematica*, the symbol = = can be used to test the truth of propositions directly:

```
5 + 2 == 6 + 1
True
```

However, testing whether the value $x = 3$ satisfies $2x^2 - 3x - 7 = 0$ yields

```
2 × 3^2 - 3 × 3 - 7 == 0
False
```

hence, $x = 3$ is not a solution of this equation. In fact, $2 \times 3^2 - 3 \times 3 - 7 = 2$, as can be seen by direct evaluation:

```
2x^2 - 3x - 7 /.x → 3
2
```

However, when a root of the equation is used:

```
Simplify [2x^2 - 3x - 7 /.x → 1/4 (3 - √65)]
0
```

FUNCTIONS

As noted earlier, an important aspect of using technology is to remember in a given context what values have been progressively assigned (or cleared and reassigned) to various constants such as a, b, c … rules defined (or cleared and redefined) for functions such as $f(x)$, $g(x)$ … and the like. A common error is to assign a constant value to the letter a, for example, and then proceed with analysis where a is intended to be an *undetermined* parameter in the rule of a function. In other cases, the '=' symbol may be used to define special relationships within one context, for example $a = b + c$, which are then not cleared for work in another context.

Different technologies, including CAS, have various ways of dealing with these issues, and the technology user needs to be familiar with these and apply the appropriate protocols for that technology to ensure that the analysis they *think* they are conducting is, in fact, what they intend. For example, if the rule of a linear function is defined by:

```
f[x_] := ax + b
```

Then a and b are undetermined coefficients:

```
Solve[f[x] == c, x]
{{x → -(b - c)/a}}
```

If, however *a* and *b* are assigned particular values:

```
{a = 2, b = 3}
```

then these will be used in future computations:

```
f[4]
11
```

And, if not cleared, will still be in use, and may lead to unexpected results:

```
Expand[(a + b)(a - b)]
-5
Clear[a,b]
Expand[(a + b)(a - b)]
a^2 - b^2
```

One of the first mathematical functions that students meet is the '*counting on*' or *successor* function, *s*, on the set of natural numbers, *N*, and is central to the axiomatic model for the natural numbers developed by Peano in 1889 and links these to set theory (see Enderton 1977 for a detailed discussion on how *number* can be developed from notions of *set* and *function*). This function has the rule $s(n) = n + 1$, and its application in *N* can be summarised diagrammatically as: $0 \to 1 \to 2 \to 3 \to \ldots$ In fact the set of natural numbers, *N*, is generated by the repeated applications of the successor function to the initial element 0:

$$0$$
$$s(0) = 1$$
$$s(s(0)) = s(1) = 2$$
$$s(s(s(0))) = s(2) = 3$$
.
.
.

Arithmetic operations are also examples of functions that students meet early on; indeed, it is the efficient and reliable evaluation of these functions that is held in high esteem in *numeracy* work:

- the *sum* function $(m, n) \to m + n$ or $+(m, n) = m + n$
- the *product* function $(m, n) \to m \times n$ or $\times(m, n) = m \times n$

Of course, in practice these functions are computed for arbitrary natural number and then integer values using the mental counting, addition and multiplication facts (tables) and written algorithms (including the combined application of place value and the distributive property for multiplication over addition) that students spend much of their early years of mathematics education becoming familiar with.

As students develop their knowledge of arithmetic operations, these are then used in various combinations to define and apply rules for more complicated functions. Students often initially meet such functions through the concept of a 'function machine' such as the following, for $x \to 2x + 3$, or $f(x) = 2x + 3$, as shown in Figures 1.12a and 1.12b:

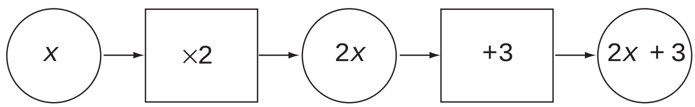


Figure 1.12a: Function machine for $x \to 2x + 3$, or $f(x) = 2x + 3$

which, after some familiarisation, is then often presented in the condensed form below, where the particular computation for $x = 3$ is shown:

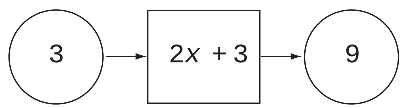


Figure 1.12b: Condensed form of function machine for $x \to 2x + 3$, or $f(x) = 2x + 3$

with an emphasis on the fact that for each 'input' (3 in this case) there is *only one* 'output' (10 in this case).

Students subsequently move in the later years of secondary schooling to the notion of function as a *reified construct*, that is, as a 'thing' which can itself be considered as an object of further manipulation and analysis. Then functions can themselves be *transformed*, *inverse* functions found (if the original function is a one-to-one function), *combined* in various ways through arithmetic operations and composition, *differentiated* and *integrated*. In short, the notion of a function as something which *acts on* other things, such as numbers, is extended to the notion of a function as something which can be *acted on* itself.

Thus, for the linear function $f(x) = 2x + 3$, with domain R, the inverse function f^{-1}, with rule $f^{-1}(x) = \dfrac{x-3}{2}$ and domain R can be found. The graphs of the f and its inverse function, f^{-1}, are reflections of each other in the line $y = x$ as shown in Figure 1.13.

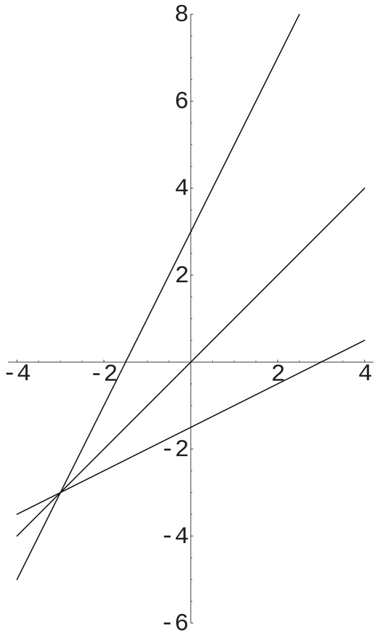


Figure 1.13: Graph of *f*, the line $y = x$ and f^{-1}

These graphs have a single point of intersection, which lies on the line $y = x$, and the coordinates of this point can be found by solving any one of the three equations:

$$f(x) = f^{-1}(x) \quad \text{or} \quad f(x) = x \quad \text{or} \quad f^{-1}(x) = x$$

For this function, the simplest equation to solve is likely to be $f(x) = x$, which gives $2x + 3 = x$ and hence $x = -3$. Since $f(x) = x$, the coordinates of this point of intersection will be $(-3, -3)$.

Similarly, other functions that are uniquely determined by *f* can also be found, such as the reciprocal function $g = \frac{1}{f}$, with the graphs of *f* and *g* shown (including a vertical asymptote for the graph of *g* at $x = \frac{-3}{2}$) in Figure 1.14.

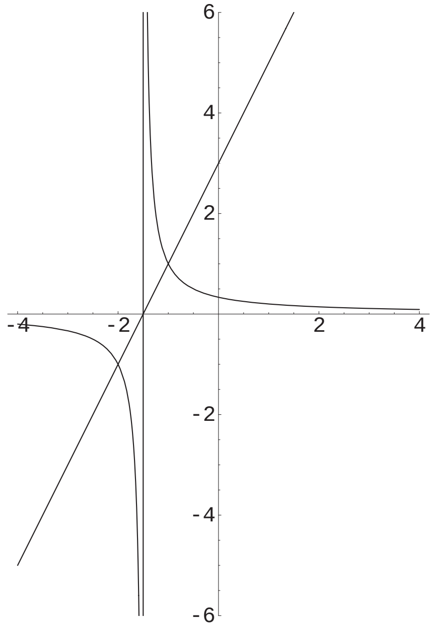


Figure 1.14: Graphs of f and $g = \frac{1}{f}$

CAS work well with functions as reified constructs, and the rules of these functions can be operated on algebraically and manipulated in various ways. For example, given g, as previously defined, $g(x)$ and $g(x-3)+2$ can be compared graphically over the interval [–4, 4], as shown in Figure 1.15.

```
f[x _]: = 2x + 3      g[x _]: = 1/f[x]
Plot[{g[x], g[x − 3] + 2}, {x, −4, 4}, PlotRange → {−5, 8}]
```

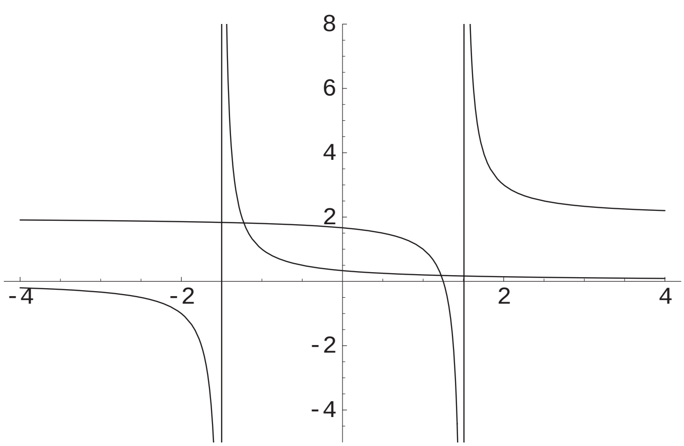


Figure 1.15: Graphs of $g(x)$ and $g(x-3)+2$ over the interval [–4, 4]

Likewise, the rule of the sum function $f + g$ can be readily obtained and subsequently used in further analysis, as required, see Figure 1.16.

```
f[x _] := 2x + 3     h[x _] := f[x] + g[x]
Plot[h[x], {x, −4, 4}, PlotRange → {−10, 10}]
```

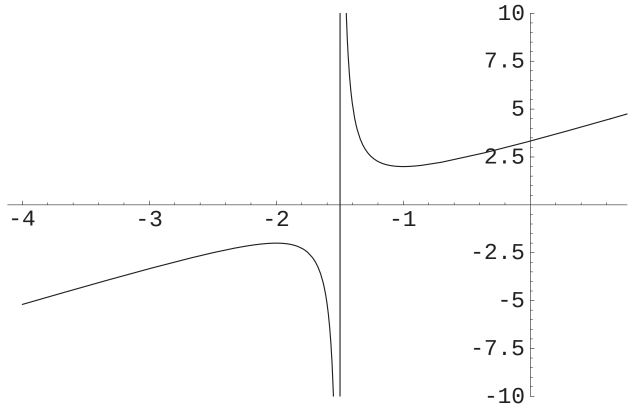


Figure 1.16: Graph of $h(x) = f(x) + g(x)$ over the interval [–4, 4]

The stationary points of the graph of the function h can be identified:

```
Solve[h ' [x] == 0, x]
{{x → -2}, {x → -1}}
```

and their corresponding function values obtained:

```
{h[−2], h[−1]}
{-2, 2}
```

Although not all CAS deal with the issue of multiple senses of '=' in such an explicit form, the underpinning considerations are relevant to all CAS and other software with similar functionality, and users of these technologies should be conversant with the relevant conventions and processes of the technology that they are using.

FUNCTIONAL EQUATIONS

So what are *functional* equations? Simply put, they are equations that involve *functions* as undetermined objects required to satisfy the given equation. That is, functions are considered as *reified constructs* that are themselves the solutions of functional equations. The solution of functional equations, in some form or other, is one of the oldest topics of mathematical analysis.

Although the systematic study of such equations is a relatively recent, and contemporary, area of mathematical study, they have been considered implicitly in various forms since antiquity and more explicitly by mathematicians such as Euler in the 18th century and Cauchy in the 19th century. Modern works on functional equations and their applications can be found increasingly from the 1960s, with applications in fields such as geometry, nomography, physics, mechanics, probability, information theory and economics. A search on the internet using *functional equations* will yield a substantial range of articles, papers and references, many involving functional equations that bear the name or names of those who first studied them as a particular topic of interest in a given theoretical or application context.

Key references from the modern era are *Lectures on functional equations and their applications* (Aczél 1966a); *On applications and theory of functional equations* (Aczél 1966b) and *Functional equations in economics* (Eichhorn 1978). A brief and very readable introduction to functional equations suitable for senior secondary mathematics can be found in Alsina (2000). The study of functional equations also provides a powerful approach to working with important concepts and relationships in function, algebra and probability, such as *symmetry*, *linearity* and *equivalence*.

Functional equations can be taken to mean almost any sort of equation involving the determination of unknown functions, and thus involve operator equations, differential equations and integral equations, as indeed is the case for many applications in the physical sciences. In this sense, the equation $f'(x) = f(x)$ is also a functional equation, since the solution sought is a *function* whose rate of change is the value of the function. The well-known general solution of this differential equation involves an arbitrary constant:

```
DSolve[f'[x] == f[x],f[x],x]
{{f[x] → e^x C[1]}}
```

however, a specific solution can be found by providing a boundary condition:

```
DSolve[{f'[x] == f[x],f[0] == 10},f[x],x]
{{f[x] → 10e^x}}
```

This is sometimes summarised for student by saying that *the* exponential function (that is, the exponential function with base e) is the function which is its own *derivative* function.

In some respects, functional equations are similar to differential equations; indeed, in some characterisations (but *not* the one which is used in this resource) they are considered to include differential equations. A more specific characterisation is to consider 'functional equations' as being those equations constructed from a finite number of unknown functions in a finite number of independent variables. In this sense, they then include difference equations, iteration equations and equations defining implicit functions, many of which have application in the humanities and in studies such as economics.

Functional equations can also be used to express algebraic structural properties, such as the commutative and associative laws. If S is a set; x, y, z are arbitrary elements of S and $\bullet$ is a binary operation on S, then the commutative and associative laws are respectively:

$$x \bullet y = y \bullet x \quad \text{and} \quad x \bullet (y \bullet z) = (x \bullet y) \bullet z$$

If a function f is defined such that $f(x, y) = x \bullet y$, then these laws have the respective functional equation forms:

$$f(x, y) = f(y, x) \quad \text{and} \quad f(x, f(y, z)) = f(f(x, y), z)$$

In this resource, only functional equations involving continuous (and generally differentiable) functions of a single real variable will be considered, with the exception of the chapter on difference equations, where discrete functions that map from N to R are considered.

This limits the field of study substantially, but provides a good basis for a useful range of investigations suitable for the senior secondary mathematics curriculum. While some calculus will be useful in parts of the discussion and development of the material, the material should be seen as *introductory* in nature, and by and large it is presented without the use of calculus and formal proofs. Indeed, it is a feature of functional equations within this more specific characterisation that the application of their theory and practice in fields like economics—for example, interest and annuity formulas—does *not* require the usual assumptions of calculus.

A solution to a functional equation is a *function* (or class of functions) so, as indicated by various examples given earlier in this chapter, consideration of both *rule* and corresponding *domain* are central. Indeed, to say that a given function is a solution to a functional equation is to assert that one can verify that it satisfies the equation for the values of the independent variable across its domain. The emphasis in the following material will be the use of functional equations to explore well-known properties in the analysis of continuous and differentiable real functions that arise in the study of functions and graphs, algebra and calculus, in senior secondary school mathematics, and some of their basic theoretical and practical applications.

The functional equations $f(x) = f(-x)$ and $f(-x) = -f(x)$ characterise those real valued continuous functions whose graphs exhibit *vertical symmetry* by reflection in the y-axis, and half-turn *rotational symmetry* about the origin, respectively; while the functional equation $f(x) = f(x + k)$ for some non-zero real constant k, is used to define *periodicity* for real valued continuous functions. The humble successor function discussed earlier is the solution to the functional equation $f(n + 1) = f(n) + 1$ where $f(0) = 0$ for domain N.

Functional equations of the form $f(x + 1) = a\,f(x) + b$, where a and b are real constants, are called first order *recurrence* relations or *difference* equations when the domain of the required solution function is the set of natural numbers, N. These include the special cases of arithmetic sequences, $f(x + 1) = f(x) + b$, where $a = 1$ and geometric sequences $f(x + 1) = a\,f(x)$, where $b = 0$. Other functional equations of this kind include the equation $f(x + 1) = (x + 1)\,f(x)$, related to the factorial function, and the equation $f(x) = f(x - 1) + f(x - 2)$, related to Fibonacci sequences.

The applications of this type of functional equation, whose solutions are *non-continuous* functions over N, in growth and decay contexts from biology and business are well known (see, for example, Chapters 4 and 10 of Hodgson & Leigh-Lancaster 1990). The *logistic* equation for population models, often used to demonstrate chaotic behaviour—depending on the parameters used in defining the function and/or the initial conditions—is a functional equation of the form $f(x) = a\,f(x - 1)\,(1 - f(x - 1))$, with solution sought over the domain N. This type of functional equation is often included in discrete mathematics-based senior secondary mathematics courses.

If real-valued differentiable functions (with continuous derivatives) are considered, functional equations that relate $f(x + y)$, $f(x - y)$, $f(xy)$ and $f(\frac{x}{y})$ to $f(x)$ and $f(y)$ can be used to characterise properties of particular functions, for example, the functional equation $f(xy) = f(x)\,f(y)$ is satisfied by power functions with rules of the form $f_q(x) = x^q$, where q is a rational constant, over their natural domain. For example, if $q = 2$, the functional equation $f(xy) = f(x)\,f(y)$ expresses the relationship $(xy)^2 = x^2y^2$, over the natural domain R, while if $q = \frac{1}{2}$, the functional equation $f(xy) = f(x)\,f(y)$ expresses the relationship $\sqrt{xy} = \sqrt{x}\sqrt{y}$ over the natural domain $R^+ \cup \{0\}$. Similarly, the logarithm function is a well-known solution of the functional equation $f(xy) = f(x) + f(y)$ over the natural domain R^+ (see also Chapters 14 and 16, Binmore 1977).

SUMMARY

Functional equations are equations whose solutions are functions. The solution of a functional equation (or the rejection of a function as a possible solution to a functional equation) may involve computation, tables, graphs and algebra.

- Equality '=' is an equivalence relation as it is *reflexive* ($x = x$); *symmetric* ($x = y$ implies $y = x$) and *transitive* ($x = y$ and $y = z$ implies $x = z$).
- Less than '<' and less than or equal to '≤' are *not* equivalence relations.
- 'Polynomial re-expression' is an equivalence relation.
- '=' is used in three ways: to *assign a fixed value* to a *constant*, for example, $a = 2$; to specify an equation which is *satisfied* by *particular* value(s) of a variable, for example $2x + 3 = 7$; and to specify the *rule* of a function, in which the variable can assume any value in the domain of the function, for example, $f: R \to R, f(x) = x^2$.
- An equation may have: *no* solutions; a *finite* number of solutions; or *infinitely* many solutions.
- A function is a correspondence between two sets X and Y, that assigns (maps) elements of one set, X, to elements of another set, Y, with the following property: each element in the set X is assigned (mapped) to exactly one element in the set Y.
- A function is a set of ordered pairs for which no two ordered pairs have the same first element.
- If a rule of the form $y = f(x)$ is available to describe the mapping, then the function f can be defined by $f = \{(x, y): y = f(x) \text{ and } x \in X\}$.
- The graph of a function is its representation by points on a cartesian coordinate system (two axes perpendicular to each other through a fixed reference point called the *origin*) where each ordered pair corresponds to the coordinates of a point.
- The set X is called the *domain* of f, written dom(f) or d_f, and the set Y is called the *co-domain* of f; the set of $f(x)$ values resulting from the application of $f(x)$ to all x in X is called the *range* of f, written ran(f) or r_f, and is a subset of Y, that is, $r_f \subseteq Y$.
- Functional equations can: represent algebraic identities or structural properties; represent symmetry and linearity relationships; and model theoretical and practical applications involving difference equations.

STUDENT ACTIVITY 1.1

For each of the following relations, decide which of the three conditions for an equivalence relation are satisfied, providing a counter-example in each case where a condition is not satisfied.

Assume a suitable natural domain where one is not specified. Hence decide whether the relation is an equivalence relation or not.

a ‘is older than’
b ‘$\leq$’ on R
c ‘$\subseteq$’ on subsets of the Roman alphabet
d ‘divides’ on N, where x divides y means there exist a natural number n such that $y = nx$

STUDENT ACTIVITY 1.2

For each of several different types of graphics calculator and/or CAS, investigate how they deal with:

- assigning constants
- denoting variables
- defining rules of functions
- solving equations
- clearing definitions and assignments
- specifying conditions or constraints in defining functions and/or solving equations

Summarise the similarities and differences in approaches between technologies and models.

STUDENT ACTIVITY 1.3

a Find the number which makes the equation $28 + 36 = 24 + \square$ true, and explain your reasoning.
b Find the number which makes the equation $108 - 47 = \square - 45$ true, and explain your reasoning.
c Use the two terms $n - 1$ and $n + 5$ and the numbers 1 and 7 to form an equation using only the operation of addition that is true for $n \in N$.
d For what values of n is the equation $2(n - 3) + 4 = n + n - 3$ true?

STUDENT ACTIVITY 1.4

a Determine the values of x for which $4x - 7 = 2(x - 3) + 6$
b Determine the values of x for which $4x - 7 = 2(2x - 3) - 1$
c Determine the values of x for which $4x - 7 = 4(x - 3) + 6$

STUDENT ACTIVITY 1.5

a Describe the relationship between equality of two functions and equality of their derivative functions.
b Describe the relationship between equality of two functions and equality of their anti-derivative functions.
c Describe the relationship between equality of two functions and equality of their definite integral over a given interval.

STUDENT ACTIVITY 1.6

a Show that the graph of a quadratic polynomial function is symmetrical by reflection in the vertical line that passes through its vertex.
b Show that the graph of a cubic polynomial function is symmetrical by half-turn rotation about its point of inflection.

References

Aczél, J 1966b, *On applications and theory of functional equations*, Elemente der Mathematik vom Höheren Standpunkt aus, vol. 5), Basel and Stuttgart.
Crossley, JN 1972, *What is mathematical logic?* Oxford University Press, Oxford.
Stillwell, J 1999, *Number*, Springer, New York.

Websites

http://www.cut-the-knot.org/language/index.shtml – Alex Bogomolny
A mathematical miscellany covering many topics and aspects of mathematics. Has an extensive glossary. This site contains a broad range of articles and materials, including content related to equations, equality and equivalence, variables, functions and functional equations, and recurrence relations.

http://en.wikipedia.org/wiki/Mathematics – Wikimedia Foundation
A comprehensive free online mathematics encyclopaedia. See, in particular, references related to equations, equality and equivalence, variables, functions and functional equations, and recurrence relations.

http://www.iscid.org/encyclopedia/Equivalence_Relation – International Society for Complexity, Information, Design
Cross disciplinary forum, which provides an encyclopaedia of science and philosophy; in particular incorporates discussion of concepts.

CHAPTER 2

AN INTRODUCTION TO FUNCTIONAL EQUATIONS

TWO SIMPLE FUNCTIONAL EQUATIONS

The study of functional equations is a contemporary area of mathematics that provides a powerful approach to working with important concepts and relationships in function and algebra such as *symmetry*, *linearity* and *equivalence*. Although the systematic study of such equations is a relatively recent area of mathematical study, they have been considered earlier in various forms by mathematicians such as Euler in the 18th century and Cauchy in the 19th century.

The study of simple functional equations can be used to provide a general framework for conceptualising and understanding key aspects of *function* and *algebra* in the senior secondary mathematics curriculum, and can also enable students to avoid some of the pitfalls commonly associated with work in these related areas of study, in particular *algebraic equivalence* (identity).

Two simple examples of functional equations are:

$$f(x) = f(-x) \qquad \textit{Functional equation } (1)$$

and

$$f(x + y) = f(x) + f(y) \qquad \textit{Functional equation } (2)$$

A solution to a functional equation is any function, f, that satisfies the equation for all values, or combinations of values, of the variable from its natural domain (that is, the largest set of values for which the function is defined). In the examples and problems considered here, these functions will generally be differentiable functions of a *real* variable.

While a function, f, is typically defined by specification of both its *domain*, dom $f = d_f$, and its *rule*, $f(x) = \ldots$, it is sometimes referred to implicitly by its rule $f(x)$, as the corresponding *natural* domain (or, as it is sometime called, the *implied* or *maximal* domain) is *assumed*. For example, the function $f: R \rightarrow R, f(x) = x^2$ might be referred to as $f(x) = x^2$ with the natural domain of R *understood*.

The following discussion can be used to introduce students to the notion of functional equations. At some stages in the implementation of a mathematics curriculum it will be important to refer to functional equations explicitly, whereas at other times they can be considered as recurring *theme* underpinning investigation of the algebraic and graphical behaviour of functions throughout a course, and it is expected that teachers will move between these approaches as appropriate. Work on functional equations also highlights two important principles of mathematical reasoning:

- *proof* of a *general* mathematical statement typically involves algebraic reasoning using an *arbitrary* or *free* variable, which is subsequently universally quantified
- *disproof* of a general mathematical statement typically involves identification of a *single counter-example*, that is, an object which satisfies the conditions of the general mathematical statement but not its conclusion

Both these aspects of mathematical reasoning are highlighted in the following discussion. The functional equations (1) and (2) have been chosen as introductory examples to illustrate the related analysis as they:

- are accessible and familiar with respect to basic background knowledge and skills
- have simple solutions
- provide an example involving one variable and another example involving two variables
- provide a context for discussion of two important concepts in mathematics—*symmetry* and *linearity*

SOLUTIONS TO THESE FUNCTIONAL EQUATIONS

EXAMPLE 2.1

Consider the functional equation $f(x) = f(-x) \ldots (1)$. The function $f(x) = x^2$ is a solution to (1) since $f(-x) = (-x)^2 = -x \times -x = x^2$ and $f(x) = x^2$ by definition.

This can also be seen from a graphical consideration of the problem. The functional equation $f(x) = f(-x)$ tells us that for any function f satisfying this equation, the negative of a given domain value has the same function output as the value itself, that is, the graph of the function will be *symmetrical* about the vertical coordinate axis, the y-axis, where the axis is the mirror line for reflection.

This is a well-known property of the graph of $f(x) = x^2$, as shown in Figure 2.1.

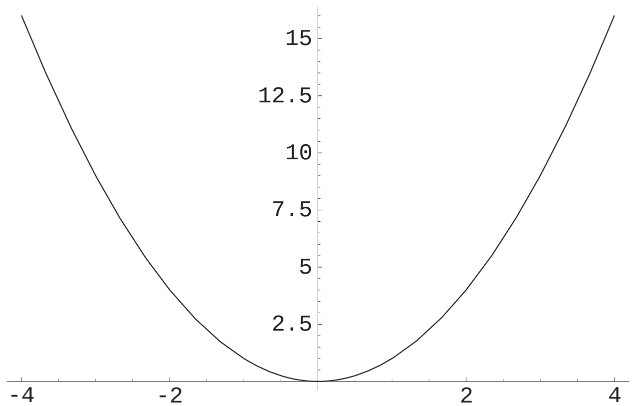


Figure 2.1: Graph of $f(x) = x^2$

While a table of values such as Table 2.1 below can be used to gain an idea of the likely relation, or for illustrative purposes, it is not conclusive, as evaluations of expression for only a finite subset of (typically integer) values from the natural domain a function can be represented.

Table 2.1: values of *f*(*x*) and *f*(–*x*) for integer values of *x* from 0 to 10

```
f[x_]:=x^2
Table[x,f[x],f[-x]}, {x,0,10}]//TableForm
```

0	0	0
1	1	1
2	4	4
3	9	9
4	16	16
5	25	25
6	36	36
7	49	49
8	64	64
9	81	81
10	100	100

This function is clearly not the only solution to (1) as can be seen from the graph of $g: R \rightarrow R$, $g(x) = \cos(x)$ shown in Figure 2.2, which also exhibits symmetry by reflection in the vertical axis.

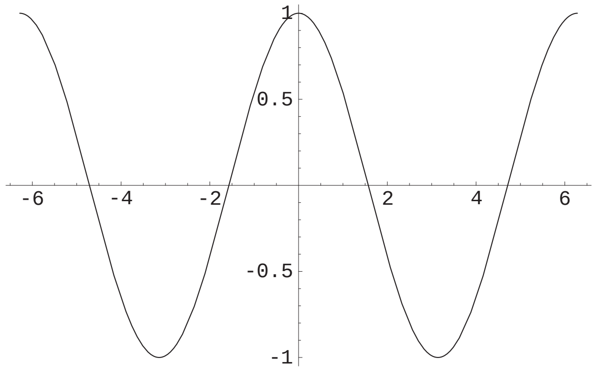


Figure 2.2: Graph of $g(x) = \cos(x)$

This can also be argued from consideration of a unit circle diagram, as shown in Figure 2.3, where the same horizontal distance, $\cos(x)$, corresponds to either the projection of the endpoint of the arc length of x units in the positive direction (anticlockwise) mapped around the circumference of the unit circle onto the horizontal axis, or the projection of the endpoint of the arc length of x units in the negative direction (clockwise) mapped around the circumference of the unit circle onto the horizontal axis.

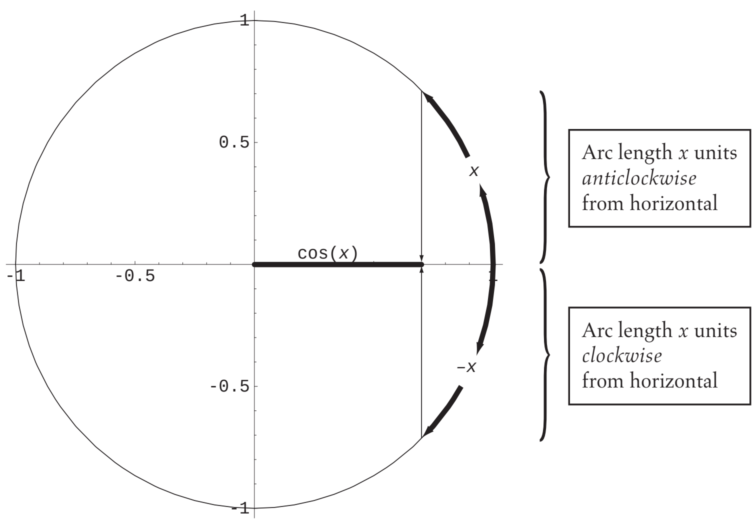


Figure 2.3: Unit circle representation that $\cos(x) = \cos(-x)$

Similarly, the absolute value, or modulus, function is a solution to (1), as can be seen from the graph of its function $h: R \to R$, $h(x) = |x|$ shown in Figure 2.4 below:

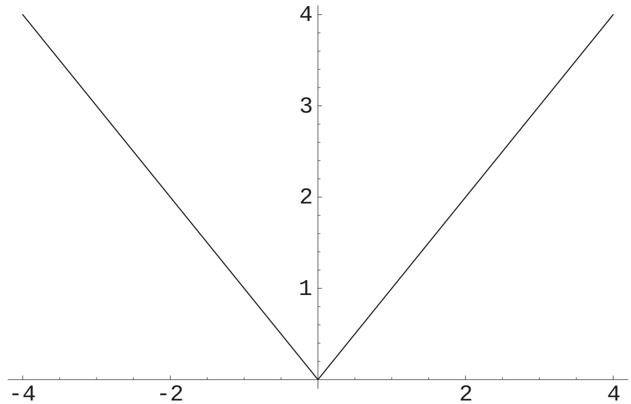


Figure 2.4: Graph of $h(x) = |x|$

In fact it is a consequence of the *definition* of the absolute value function that $|x| = |-x|$. In each of these cases, an argument has been used to show how the result applies for an *arbitrary* x, and hence applies for all x in the relevant domain of interpretation.

In many cases where the *graph* of a function is indicated, this can only actually be referring to an illustrative *part* of the graph, since either the natural domain, or the domain used to specify a function, will in many cases be an unbounded infinite set such as R. This assumes confidence that the overall behaviour of a function has essentially been captured in the illustrative part of the graph drawn. This is a subtle but important point that needs to be drawn to the attention of students, especially when technology is used to assist graphical analysis in mathematical inquiry.

Technologies such as graphics calculators and CAS provide various mechanisms for controlling the *subsets* of the *domain* and *range* of a function used in plotting a graph, given its rule, as well as applying their own built in default settings and/or procedures for optimising the key features of the function to be exhibited graphically. For example, the CAS *Mathematica* requires the rule of the function, `f[x]` to be used in conjunction with the interval `[a,b]` over which the graph is to be plotted, by execution of the command `Plot[f[x],{x, a, b}]` The vertical axis scale, and the starting point on this scale where the graph is plotted, is determined by an internal procedure. This procedure can be over-ridden by use of options for **`Plot`** such as **`PlotRange`** and **`AxesOrigin`**. Students should be encouraged to practise varying the dimensions of the part of the graph exhibited using the relevant specification procedures for the technology they are using. This will avoid their confounding the horizontal and vertical 'plot-window' or 'frame' used by a technology with the domain and range proper of a given function.

The function $f(x) = x^2$ is, however, *not* a solution to functional equation (2) since:

$$f(x + y) = (x + y)^2 = x^2 + 2xy + y^2 = f(x) + 2xy + f(y)$$

and, in general, this is *not* equal to $f(x) + f(y) = x^2 + y^2$.

The false belief that $f(x + y) = f(x) + f(y)$ for $f(x) = x^2$ is one of the most common algebraic misconceptions of students. For example, many students believe (incorrectly) that the expansion of $(x + 2)^2$ is $x^2 + 4$ rather than $x^2 + 4x + 4$. Now, it may be simply that these students 'forgot' the middle term, but whatever the reason, students who repeatedly make this error have most likely adopted a belief in a schema for expansion that leads in this case to $f(x + 2) = f(x) + f(2) = x^2 + 4$. Certainly they all too often write this—or variations of it—in their own working.

To overcome such fallacious reasoning, students require their attention to be drawn to it explicitly, and to the correct reasoning as well. The latter can be demonstrated by the use of physical manipulatives such as Algebra Experience Materials (AEM) leading to consideration of a diagram such as that shown in Figure 2.5, highlighting the shaded areas corresponding to x^2 and y^2 as a subset of the total area of $(x + y)^2$:

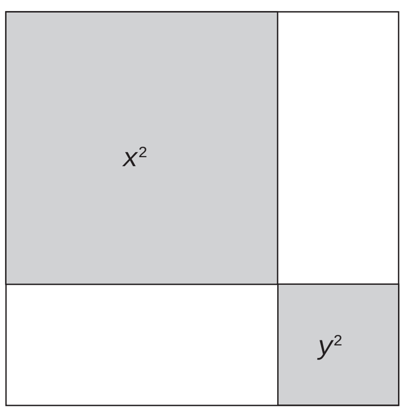


Figure 2.5: Areas of x^2 and y^2 as a subset of the total area of $(x + y)^2$

Thus the total area representing $(x + y)^2$ is comprised of x^2 and y^2 and two lots of xy. In the case where $y = 2$, this gives $(x + y)^2 = x^2 + 4 + 4x$.

For some students it will be important to exhibit a selection of counter-examples to the linearity assumption of $f(x + y) = f(x) + f(y)$ for $f(x) = x^2$. Technology can be used to test a range of combinations of values for x and y:

```
f[x_] := x^2
{f[1 + 2], f[1] + f[2]}
{9,5}
{f[-3 + 7],f[-3] + f[7]}
{16,58}
```

Indeed, a table of values such as Table 2.2, can help to illustrate *what* the difference is in this case, for example, where $y = 3$ and x varies in integer steps from –5 to 5.

Table 2.2: Computation of *f*(*x* + *y*) and *f*(*x*) + *f*(*y*)

```
Table[{x,3,f[x+3],
f[x]+f[3], f[x+3]-(f[x]+f[3])}, {x,-5,5}]//TableForm
```

-5	3	4	34	-30
-4	3	1	25	-24
-3	3	0	18	-18
-2	3	1	13	-12
-1	3	4	10	-6
0	3	9	9	0
1	3	16	10	6
2	3	25	13	12
3	3	36	18	18
4	3	49	25	24
5	3	64	34	30

It is important to note that solutions to functional equations are those functions that satisfy the equation *in general* for values of the dependent variable from the natural domain of the function.

It is possible, however, to inquire whether there are *particular* values, or combinations of values, of the independent variable for which a given equation is satisfied. Thus, it will also be important to show students that there are some *particular* combinations of values of x and y for which

$f(x + y) = f(x) + f(y)$ given $f(x) = x^2$. In fact, in this case there are *infinitely many* such values.

If $x = 0$ and y is any real number, or if $y = 0$ and x is any real number, then for such combinations of value of the independent variables x and y:

$$f(x + 0) = f(x) \quad \text{and} \quad f(x) + f(0) = f(x) + 0 = f(x)$$

and similarly:

$$f(0 + y) = f(y) \quad \text{and} \quad f(0) + f(y) = 0 + f(y) = f(y)$$

So, for any combination where either or both of x and y is zero, $(x + y)^2$ will have the same value as $x^2 + y^2$.

EXAMPLE 2.2

Consider the functional equation $f(x + y) = f(x) + f(y)$ … (2). The function h: $R \rightarrow R$, $h(x) = 2x$ *is* a solution of (2) since:

$$h(x + y) = 2(x + y) = 2x + 2y = h(x) + h(y)$$

This is, in fact, a simple application of the distributive property for multiplication over addition $a(b + c) = ab + ac$ for real numbers applied to linear expressions. However, the function h is *not* a solution of (1) since:

$$h(-x) = 2(-x) = -2x \text{ which is not the same as } h(x) = 2x$$

A straightforward counter-example is provided by any non-zero value of x, for example, if $x = 6$, then $2 \times -6 = -12$, but $2 \times 6 = 12$. This can also be observed by noting that the graph of $h(x)$ shown in Figure 2.6 is not symmetrical under reflection in the y-axis.

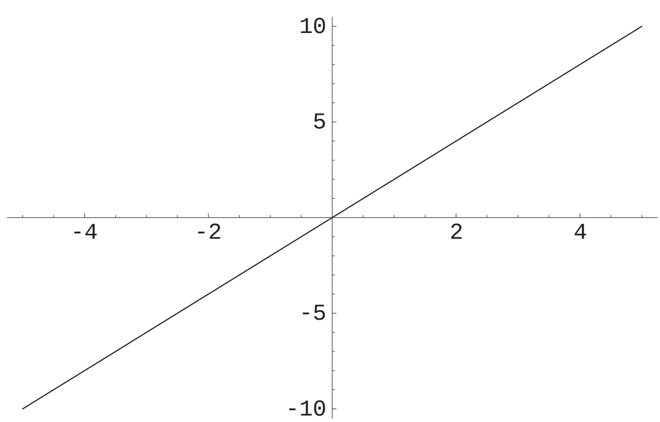


Figure 2.6: Graph of $h(x) = 2x$

While the graph of the function h is *not* symmetrical by *reflection* in the vertical axis, it does have a half-turn, or *rotational*, symmetry under the transformation R_π about the origin (0, 0). For this function, and other functions that have this symmetry, the function value for the negative of an element of the domain of the function is the negative of the function value of that element of the domain, or more concisely, $f(-x) = -f(x)$.

This relationship can be expressed as another functional equation, one which characterises a half-turn rotational symmetry about the origin for the graphs of certain functions:

$$f(-x) = -f(x) \qquad \textit{Functional equation } (3)$$

Clearly, the function $f(x) = x^2$ will *not* be a solution to (3), since it has previously been shown that it is a solution to (1).

In general, several key questions can be formulated with respect to functions and functional equations:

- For a given functional equation, what functions are solutions of that equation?
- For a given function, what functional equations does it satisfy?
- What values of the domain of a function satisfy the relationship defined by a given functional equation?

Where functional equations arise in modelling contexts, such as the spread of a disease, expectation and variance, or the construction of tax scales, information related to the situation may provide some insight into the likely nature of the solution function.

SUMMARY

The solution of functional equations involve writing equations in terms of an arbitrary function f, and then determining those functions that satisfy the equation over their natural domain:

- A function may be shown to not be a solution to a given functional equation by numerical or graphical counter-example, or by algebraic reasoning
- In general, an algebraic argument is required to show that a given function *is* a solution to a function equation for all values (or combination of values) of the independent variable over its natural domain.

- For a given function, there may be *some* values of its natural domain that satisfy the relationship defined by a particular functional equation. These values are *not* suitable as *counter-examples*.
- $f(x) = f(-x)$ is a functional equation, for which $f(x) = x^2$ is a solution, but for which $h(x) = 2x$ is not a solution.
- $f(-x) = -f(x)$ is a functional equation, for which $h(x) = 2x$ is a solution, but for which $f(x) = x^2$ is not a solution.
- $f(x + y) = f(x) + f(y)$ is a functional equation, for which $h(x) = 2x$ is a solution, but for which $f(x) = x^2$ is not a solution.

STUDENT ACTIVITY 2.1

For each of the following function rules, with their corresponding natural domain, describe, as applicable, the appearance and behaviour of the graph of the function for:

- small and large intervals centred around the origin
- large negative values of the independent variable
- large positive values of the independent variable

a $f(x) = 4x$
b $f(x) = x^2$
c $f(x) = x^3$
d $f(x) = x^4$
e $f(x) = 10$
f $f(x) = \sqrt{x}$
g $f(x) = \sin(x)$
h $f(x) = \cos(x)$
i $f(x) = \frac{1}{x}$
j $f(x) = \frac{1}{x^2}$
k $f(x) = 2^x$
l $f(x) = \log_{10}(x)$
m $f(x) = |x|$

STUDENT ACTIVITY 2.2

Identify which of the functions listed in Activity 2.1 are solutions to the functional equation $f(x) = f(-x)$. Provide either an algebraic argument or a diagram to support your decision, and check this graphically.

STUDENT ACTIVITY 2.3

Investigate which functions satisfy the functional equation $f(-x) = -f(x)$. Provide either an algebraic argument or a diagram to support your decision, and check this graphically.

STUDENT ACTIVITY 2.4

Identify any functions that are solutions to functional equations (1) and (3).

STUDENT ACTIVITY 2.5

A function f is said to be *self-inverse* if $f(f(x)) = x$ for all x in its natural domain.

a Investigate whether any of the functions used for examples in this chapter are self-inverses.

b Identify some continuous functions of a real variable that are *self-inverse*.

STUDENT ACTIVITY 2.6

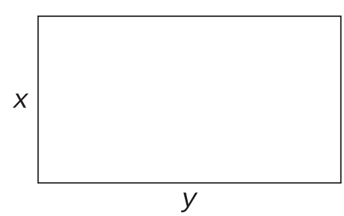
Consider the functional equation $f(x \times y) = f(x) + f(y)$.

a If $f(x) = x$, determine the values of x and y for which this functional equation is true.

b If $f(x) = x^2$, determine the values of x and y for which this functional equation is true.

STUDENT ACTIVITY 2.7

Consider the rectangle as shown in the diagram below with side lengths x and y.



Interpret each of the following functional equations with respect to the rectangle:

a $f(x, x) = 1$

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

c $f(ax, ay) = a^2 f(x, y)$

References

Alsina, C 2000, 'Mathematical modelling by means of functional equations: The missing link in the learning of functions', in *Modelling and mathematics education ICTMA 9: Applications in science and technology*, JF Matos et alia (eds), Horwood Publishing, Westergate.

Websites

http://mathworld.wolfram.com/ – Wolfram Research
Online encyclopaedia by the developers of *Mathematica*. An extensive and rigorous mathematical encyclopaedia, with extensive cross-references. See sections on functional equations, recurrence relations and sequences.

http://encyclopedia.thefreedictionary.com/Functional+equations – Farlex
A broad-ranging free dictionary/encyclopaedia across domains. Provides a good concise introduction under its *Functional equation* entry.

http://www.math.ust.hk/excalibur/v8_n1.pdf – Department of Mathematics, Hong-Kong University of Science and Technology.

Article on 'Functional Equations' in *The Mathematical Excalibur*, February–March 2003 relating to questions from the Mathematics Olympiad. Shows how solutions to some simple functional equations can be developed from *N* to *R*.

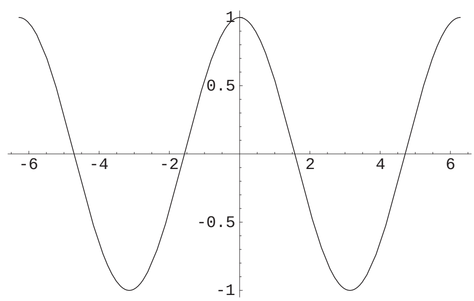
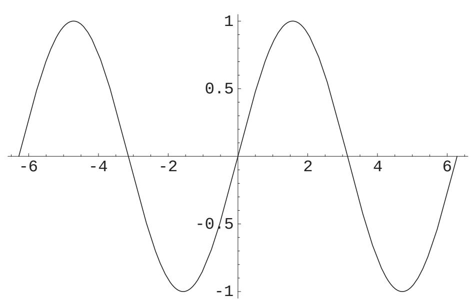
CHAPTER 3

FUNCTIONAL EQUATIONS INVOLVING *f*(*x*) AND CONSTANTS

Probably the simplest functional equations are those expressed in terms of $f(x)$ and some constants, and these are the focus of the discussion in this chapter.

Two functional equations of this form have already been considered in Chapter 1, and are summarised in Table 3.1. They characterise important symmetry properties of the graphs of certain functions.

Table 3.1: Summary of two simple functional equations

Functional equation	*Sample solution*	*Graph*
$f(x) = f(-x)$ or $f(x) - f(-x) = 0$	$f(x) = \cos(x)$	
$-f(x) = f(-x)$ or $f(x) + f(-x) = 0$	$f(x) = \sin(x)$	

These functional equations have many solutions—those functions whose graphs are symmetrical about the vertical axis by reflection in this axis as a mirror line, and those functions whose graphs are symmetrical about the origin by a half-turn rotation, respectively. Similarly, any power function of the form $f(x) = x^n$ where n is an *even* integer is a solution to the functional equation $f(x) = f(-x)$, while any power function of the form $f(x) = x^n$ where n is an *odd* integer is a solution to the functional equation $-f(x) = f(-x)$ or $f(x) = -f(-x)$.

Some texts refer to any function whose graph has the property $f(x) = f(-x)$ as an *even function* and any function whose graph has the property $f(x) = -f(-x)$ as an *odd function*.

The matrix for reflection in the y-axis (that is, the line $x = 0$) is specified by the relation

$M_y(x, y) = (-x, y)$ so:

$$M_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}$$

as shown in Figure 3.1:

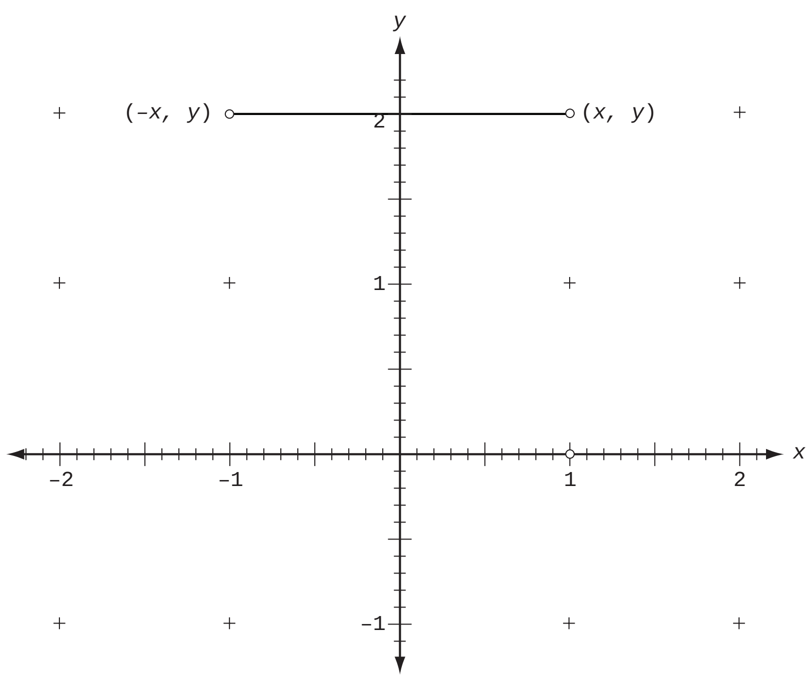


Figure 3.1: Reflection of a point in the vertical axis, $x = 0$

The effect of this transformation is to map the graph of $y = f(x)$ onto the graph of $y = f(-x)$. For functions whose graphs are symmetrical by reflection about the y-axis, this means that $y = f(x)$ and $y = f(-x)$ are identical, that is $f(x) = f(-x)$.

The matrix for a half-turn rotation anticlockwise about the origin (0, 0) is specified by the relation $R_\pi (x, y) = (-x, -y)$ so:

$$R_\pi \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}$$

This is equivalent to a reflection in the vertical axis $x = 0$ followed by reflection in the horizontal axis $y = 0$, that is $R_\pi = M_x M_y$ as can be seen from Figure 3.2:

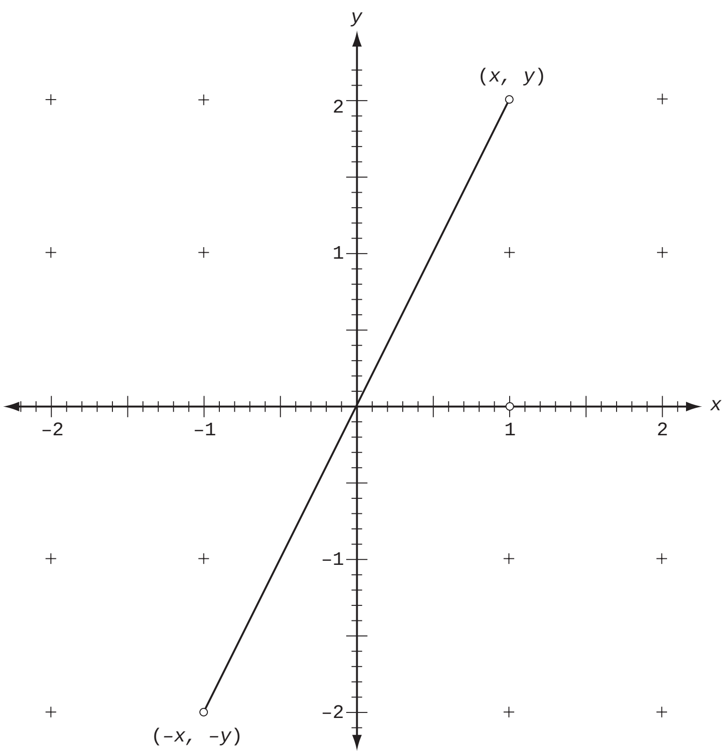


Figure 3.2: Half-turn rotation of a point anticlockwise about the origin

The effect of this transformation is to map the graph of $y = f(x)$ onto the graph of $-y = f(-x)$ or $y = -f(-x)$. For functions whose graphs are symmetrical by half-turn rotation about the origin (0, 0) this means that $y = f(x)$ and $y = -f(-x)$ are identical, that is $f(x) = -f(-x)$.

The question then arises: do functional equations of this sort ever have unique solutions or families of solutions? This question is explored further in the following discussion.

FUNCTIONAL EQUATIONS AND SCALE

EXAMPLE 3.1: THE FUNCTIONAL EQUATION $f(kx) = kf(x)$

How does one go about finding a solution, or solutions, to a functional equation such as $f(kx) = kf(x)$? A helpful starting point is to try to see if there is a simple way to interpret the functional equation. For example, as $k \neq 1$ is a fixed non-zero real constant (the solution is trivial if $k = 0$ or $k = 1$) one can assume a particular value, such as $k = 2$, and see what happens with respect to some simple known functions.

For any solution to the particular functional equation $f(2x) = 2f(x)$, the value of the function at twice the value of a given input is twice the value of the function at that given input, so, if $x = 3$, then $f(6) = 2f(3)$, and similarly for *all* values of the variable x. Now, various potential solution functions can be 'tried' empirically by inspection of tables of values; however, this process can only identify counter-examples for various functions, or, if no counter-examples are found after some systematic variation of the range of independent variable values in tables, it can *suggest* that a given function is a possible solution.

Suppose $f(x) = x^2$ and let x take integer values from –5 to 5. Table 3.2 provides a selective comparison of values of $f(2x)$ and $2f(x)$ for this function.

Table 3.2: Comparison of $f(2x)$ and $2f(x)$ for $f(x) = x^2$ for integer values of x from –5 to 5.

```
Table[{x,f[2x], 2f[x]},{x,-5,5}]//TableForm
```

-5	100	50
-4	64	32
-3	36	18
-2	16	8
-1	4	2
0	0	0
1	4	2
2	16	8
3	36	18
4	64	32
5	100	50

This table indicates three things:

- $f(x) = x^2$ is *not* a solution to the functional equation $f(2x) = 2f(x)$.
- $x = 0$ may not be a useful value to choose as a counter-example, since for this particular value of the domain the relationship is satisfied
- $f(x) = x^2$ appears to be a solution to the functional equation $f(2x) = 2 \times 2f(x) = 4f(x)$.

The third observation can be verified algebraically:

$$\text{if } f(x) = x^2 \quad \text{then} \quad f(2x) = (2x)^2 = 2x \times 2x = 4x^2 = 4f(x)$$

This also suggests a further line of investigation. Suppose $f(x) = x^n$, where n is a positive integer, then, in general $f(kx) = (kx)^n = k^n x^n$. For this to be a solution to the functional equation $f(kx) = kf(x)$ it is required that $k^n x^n = kx^n$. This will be true only when $n = 1$, so there is a solution: $f(x) = kx$.

Thus, when $k = 2$, a solution is $f(x) = 2x$, since $f(2x) = 2(2x) = 4x$ and $2\,f(x) = 2 \times 2x = 4x$. There are other solutions; indeed, any function of the form $f(x) = ax$ will be a solution to the functional equation $f(kx) = kf(x)$ as $a(kx) = k(ax)$ for any real numbers a, k and x.

On the other hand, the function $g(x) = \sin(x)$ is not a solution to the functional equation $f(2x) = 2f(x)$, since this would imply that $\sin(2x) = 2\sin(x)$. While many students appear to believe, incorrectly, that this relationship is true, the corresponding graphs, as shown in Figure 3.3, clearly indicate that this is *not* the case.

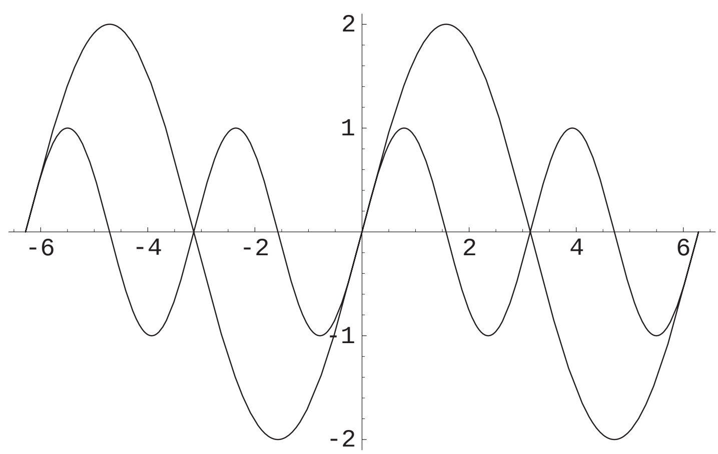


Figure 3.3: Graphs of sin(2x) and 2sin(x)

Consideration of the unit circle will indicate that the equation $\sin(2x) = 2\sin(x)$ does in fact have infinitely many solutions—one at each integer multiple of π. However, it is also clear that the graphs of the functions, and hence the values of the functions, differ across the rest of the natural domain of the sine function, R. In general, $g(x) = \sin(x)$ is not a solution to the functional equation $f(kx) = kf(x)$, for any real non-zero values of k other than $k = 1$ or $k = -1$.

The solution of the functional equation $f(kx) = kf(x)$ can also be considered in terms of transformations. The transformation of dilation by factor k from the horizontal axis in the vertical direction, $D_{1,\,k}$ transforms $y = f(x)$ into $\frac{y}{k} = f(x)$ or $y = kf(x)$. Similarly, the transformation of dilation by factor $\frac{1}{k}$ from the vertical axis in the horizontal direction, $D_{\frac{1}{k},1}$ transforms $y = f(x)$ into $y = f\left(\frac{x}{\frac{1}{k}}\right)$ or $y = f(kx)$. Thus, those functions whose graphs are identical under the transformations $D_{1,\,k}$ and $D_{\frac{1}{k},1}$ are those functions which are solutions to the functional equation $f(kx) = kf(x)$. The graphs of linear functions where $f(x) = ax$, and a is a non-zero real number, have this property.

FUNCTIONAL EQUATIONS AND PERIOD

EXAMPLE 3.2: THE FUNCTIONAL EQUATION $f(x + k) = f(x)$

For this functional equation, k is taken to be a non-zero real constant. Interpreting this functional equation tells us that for *every* value x in the domain of f, there is a corresponding value $x + k$, also in the domain of f, such that the function value at x is the same as the function value at $x + k$. A straightforward solution to this functional equation is the family of constant functions with natural domain R where: $\{f_a: f(x) = a$, and a is a real constant$\}$.

For example, if $f_{10}(x) = 10$ for all real x, then $f_{10}(x) = 10$ and $f_{10}(x + k) = 10$ for any real value of k, can be seen in Figure 3.4.

Figure 3.4: Graph of $f_{10}(x) = 10$

Some students may believe that the quadratic function $f(x) = ax^2 + bx + c$ is a possible solution because of the symmetry property $f\left(-\frac{b}{2a} - h\right) = f\left(-\frac{b}{2a} + h\right)$, for real values of h. For these students it will need to be carefully pointed out that the distance between the x values, $x = \frac{-b}{2a} - h$ and $x = \frac{-b}{2a} + h$ with the same $f(x)$ value is not a constant k, but is variable and equal to $2h$. For example, if $f(x) = x^2$ then it is true that $f(-2) = f(2)$, in which case k would need to be 4, since $f(-2 + 4) = f(2)$. However, $f(-5) = f(5)$, in which case k would need to be 10, since $f(-5 + 10) = f(5)$, but k must have a fixed value, so this function is *not* a solution to the functional equation $f(x + k) = f(x)$.

In fact, the functional equation $f(x + k) = f(x)$ provides the definition of a *periodic* function. This may also be considered in terms of transformations, since a translation by k units horizontally to the left, or $T_{-k,\,0}$ transforms $y = f(x)$ onto $y = f(x - (-k))$ or $y = f(x + k)$. Thus, any function whose graph is identical to that of the same function translated k units to the left will be a solution of the functional equation $f(x + k) = f(x)$. The function h: $R \rightarrow R$, where $h(x) = \sin(nx)$, and n is a non-zero real constant, is a well-known periodic function with this property.

Once a solution is found, then others can be generated by various transformations, where the amplitude is a multiple of the first solution, or where the phase is an integer multiple of k, or where the period is an integral fraction of the period of the first solution. An initial solution can be found by horizontal dilation of the basic sine function, with period 2π. Suppose, as is commonly the case where hourly values of a function are determined throughout a day, that the required period is 24 hours, that is, $k = 24$. Then we need to apply a horizontal dilation factor m, $D_{m,1}$ such that $2\pi m = 24$, or $m = \frac{12}{\pi}$. The dilation $D_{m,1}$ transforms $y = f(x)$ into $y = f\left(\frac{x}{m}\right)$, or in this particular case, transforms $y = \sin(x)$ into $y = \sin\left(\frac{x}{\frac{12}{\pi}}\right)$ or $y = \sin\left(\frac{\pi x}{12}\right)$. The graph of this function is shown in Figure 3.5, where it can be seen that for $h(x) = \sin\left(\frac{\pi x}{12}\right)$ we have $f(x + 24) = f(x)$.

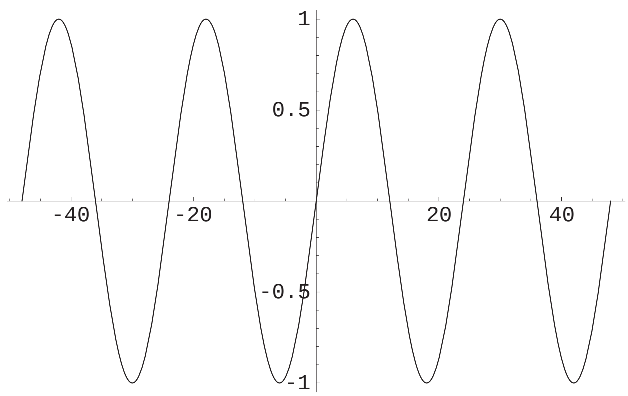


Figure 3.5: Graph of $h(x) = \sin\left(\frac{\pi x}{12}\right)$

In general, for arbitrary k, the circular function with rule $h(x) = \sin\left(\frac{2\pi x}{k}\right)$ is a solution to the functional equation $f(x + k) = f(x)$.

More generally, periodic behaviour is displayed by infinite sums of sine and cosine functions using what are called *Fourier series*. These can be used to model a range interesting waveforms such as square and triangular waves. In 1807, the French mathematician Joseph Fourier described how a rule for an arbitrary function defined over the interval $(-\pi, \pi)$ could, when certain conditions are satisfied, be represented in the form: $\sum_{n=0}^{\infty}(a_n\cos(nx) + b_n\sin(nx))$. Such expressions are called *Fourier series.*

The coefficients in the expansion of Fourier series are given by the relations: $a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx$ $\quad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx$ and $b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx$ and using graphs for functions with rules represented in the form $\sum_{n=0}^{\infty}(a_n\cos(nx) + b_n\sin(nx))$, it is possible to represent various waveforms such as square waves and triangular waves, a somewhat surprising result given the curved nature of graphs of the sine and cosine functions.

For example, square waves can be represented by the graphs of Fourier series as shown in Figures 3.6a and 3.6b.

$$S(x) = \frac{1}{2} - \frac{2}{\pi}\sum_{n=1,3,5\ldots}\left(\frac{\sin(nx)}{n^2}\right)$$

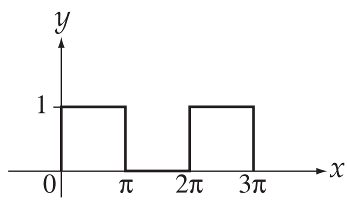


Figure 3.6a: Graph of sine-based Fourier series

and triangular waves may be represented by the Fourier series:

$$T(x) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1,3,5\ldots}\left(\frac{\cos(nx)}{n^2}\right)$$

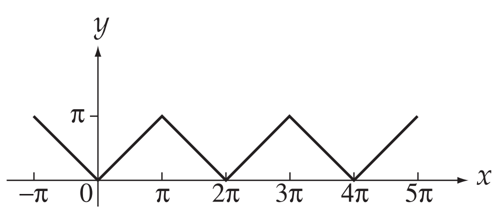


Figure 3.6b: Graph of cosine-based Fourier series

Computer algebra systems are an ideal tool for exploring waveforms based on square waves, triangular waves and combinations of these waveforms. This is a good context for some open-ended student exploration of mathematical ideas and their applications using technology. Such work could initially proceed from consideration of graphs of functions with rules of the form:

$f_1(x) = a_1\sin(x)$

$f_2(x) = a_1\sin(x) + a_2\sin(3x)$

$f_3(x) = a_1\sin(x) + a_2\sin(3x) + a_3\sin(5x)$

and so on for different combinations of values of the coefficients.

This could subsequently be extended to investigation of combinations of expressions for square and triangular waveforms to develop other Fourier series with graphs that represent waveforms such as those shown in Figure 3.7.

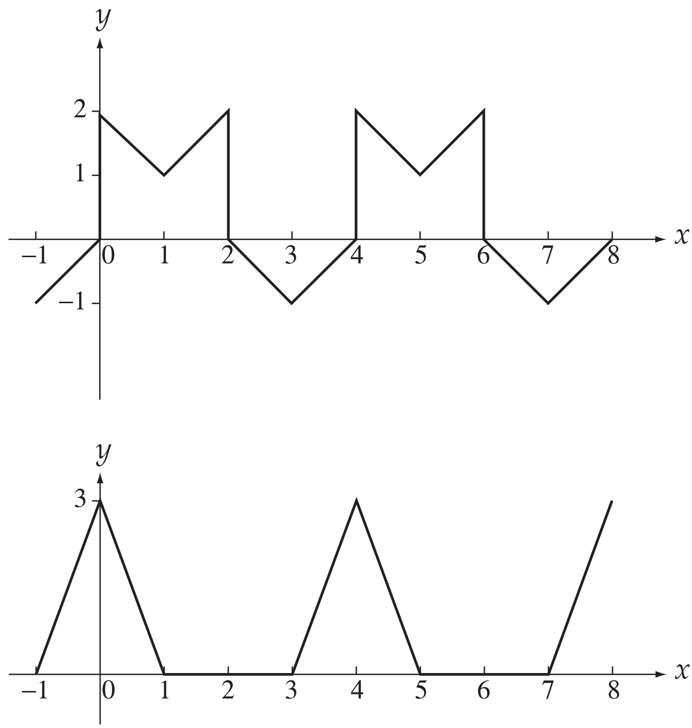


Figure 3.7: Graphs of more complicated waveforms

SUMMARY

- Functions that satisfy the functional equation $f(x) = f(-x)$ are called *even* functions and their graphs exhibit reflection symmetry in the y-axis, that is, the vertical line $x = 0$. Examples of even functions are x^2, $\cos(x)$ and $|x|$.
- Functions that satisfy the functional equation $-f(x) = f(-x)$ or $f(x) = -f(-x)$ are called *odd* functions and their graphs exhibit half-turn rotational symmetry about the origin (0, 0). Examples of odd functions are x^3, $\sin(x)$ and $\frac{1}{x}$.
- Many students mistakenly assume, or believe, that the functional equation $f(kx) = kf(x)$ is satisfied by almost all functions and hence often make 'algebraic' errors such as the following in their working:
  - $\sin(2x) = 2\sin(x)$
  - $(3x)^4 = 3x^4$
  - $\log(5x) = 5\log(x)$

  Explicit attention to this functional equation should provide students with the understanding that will enable them avoid these errors.
- The functional equation $f(x + k) = f(x)$ describes *periodic* functions, and a solution to this equation is the circular function of the form $h(x) = \sin\left(\frac{2\pi x}{k}\right)$.

STUDENT ACTIVITY 3.1

a Show that the function $f(x) = 2^x$ is neither odd nor even, and draw the graphs of $f(x)$, $f(-x)$, $-f(x)$ and $-f(-x)$ on the same set of axes.

b Show that the *relation* $y^2 = x^2$ is both even and odd, and draw its graph.

STUDENT ACTIVITY 3.2

Use tables, graphs or algebraic reasoning to decide whether each of the following functions is a solution to the functional equation $f(kx) = kf(x)$.

a $f(x) = 4x$
b $f(x) = x^2$
c $f(x) = x^3$
d $f(x) = x^4$
e $f(x) = 3$

STUDENT ACTIVITY 3.2 (CONTINUED)

f $f(x) = \sqrt{x}$
g $f(x) = \sin(x)$
h $f(x) = \cos(x)$
i $f(x) = \frac{1}{x}$
j $f(x) = \frac{1}{x^2}$
k $f(x) = 2^x$
l $f(x) = \log_{10}(x)$
m $f(x) = |x|$

STUDENT ACTIVITY 3.3

Use tables, graphs or algebraic reasoning to decide whether each of the following functions is a solution to the functional equation $f(x + k) = f(x)$.

a $f(x) = 4x$
b $f(x) = x^2$
c $f(x) = x^3$
d $f(x) = x^4$
e $f(x) = 3$
f $f(x) = \sqrt{x}$
g $f(x) = \sin(x)$
h $f(x) = \cos(x)$
i $f(x) = \frac{1}{x}$
j $f(x) = \frac{1}{x^2}$
k $f(x) = 2^x$
l $f(x) = \log_{10}(x)$
m $f(x) = |x|$

STUDENT ACTIVITY 3.4

Investigate solutions of the following functional equations.

a $f(x) = \frac{1}{f(x)}$
b $f^{-1}(x) = f(x)$
c $f\left(\frac{1}{x}\right) = \frac{1}{f(x)}$

STUDENT ACTIVITY 3.5

Discuss the relationship between the functional equation $f(-x) = -f(x)$, the functional equation $f(kx) = kf(x)$ and their solutions.

STUDENT ACTIVITY 3.6

a Find all constant, linear and quadratic functions that are solutions to the functional equation: $f(x)\,f(-x) = f(x^2)$.

b Find a real valued differentiable function f which is a solution to the functional equation: $x^2\,f(x) + f(1 - x) = 2x - x^4$.
(Hint: replace x by $1 - x$ and compare the resultant equation with the given equation).

References

Estep, D 2002, *Practical analysis in one variable*, Springer, New York.

Websites

http://www.purplemath.com/modules/symmetry.htm – PurpleMath
This site provides an introduction to symmetry of graphs of functions and relations using straightforward examples.

http://www.teachers.ash.org.au/mikemath/resources/periodic.html – Mike Shepperd
This website contains a broad range of contexts for periodic behaviour and modelling involving circular (trigonometric) functions.

www.math-cs.cmsu.edu/~mjms/2002.3/joewiener.ps – Department of Mathematics and Computer Science, Central Missouri State University
This is a paper on functions that are involutions (self-inverses).

CHAPTER 4

FUNCTIONAL EQUATIONS INVOLVING $f(x)$, $f(y)$ AND EQUIVALENCES

One of the most common incorrect assumptions students make about many functions of a single real variable is that they satisfy the functional equation $f(x + y) = f(x) + f(y)$. Indeed, one of the major tasks of teachers is to ensure that, as much as possible, students do *not* incorporate *incorrect* 'steps' such as the following in their work on function and algebra:

$(x + y)^2 = x^2 + y^2$ $\qquad$ $\dfrac{1}{x + y} = \dfrac{1}{x} + \dfrac{1}{y}$

$\log(x + y) = \log(x) + \log(y)$ $\qquad$ $\sin(x + y) = \sin(x) + \sin(y)$

$\sqrt{x + y} = \sqrt{x} + \sqrt{y}$ $\qquad$ $2^{x + y} = 2^x + 2^y$

The functional equation $f(x + y) = f(x) + f(y)$ is called the *Cauchy* equation as it was studied in the 19th century by the famous French mathematician Augustin Louis Cauchy, who also introduced the function representation of the limit definition of the derivative in calculus:

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

The Cauchy functional equation also characterises a *linearity property* for those functions where that the *functional* value of a *sum* is the *sum* of the *functional* values. This is a rather special property that does *not* hold, in general, for most functions—contrary to the popular belief of many students.

However, it does relate to the distributive property for real numbers, and hence common algorithms for multiplication. That is, if f: $R \to R$, $f(x) = ax$ then $f(x + y) = f(x) + f(y)$ since $a(x + y) = ax + ay$. For example:

$$6 \times 17 = 6 \times (10 + 7) = (6 \times 10) + (6 \times 7) = 60 + 42 = 102.$$

Thus it is perhaps not surprising that students would have an inherent belief in its more widespread 'applicability' than is actually the case.

The linearity property *does* apply in other contexts:

- differentiation of a real-valued functions of a single variable
  $\boldsymbol{D}[f + g] = \boldsymbol{D}[f] + \boldsymbol{D}[g]$, for example
  $$\boldsymbol{D}\left[x^2 + \frac{1}{x}\right] = \boldsymbol{D}[x^2] + \boldsymbol{D}\left[\frac{1}{x}\right] = 2x - \frac{1}{x^2}$$
- integration of a real-valued function of a single variable
  $\int(f + g)\,dx = \int f\,dx + \int g\,dx$, for example
  $\int(\cos(x) + e^{2x})\,dx = \int\cos(x)\,dx + \int e^{2x}\,dx = \sin(x) + \frac{1}{2}e^{2x} + c$ where $c \in R$
- to linear (affine) transformations of vectors where $T(\boldsymbol{u} + \boldsymbol{v}) = T\boldsymbol{u} + T\boldsymbol{v}$, for example:

  if $T = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $\boldsymbol{u} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ then

  $$T(\boldsymbol{u} + \boldsymbol{v}) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 13 \end{bmatrix} \text{ and } T\boldsymbol{u} + T\boldsymbol{v} = \begin{bmatrix} 6 \\ 12 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 13 \end{bmatrix}$$
- expectation and variance of the sum of two independent random variables
  $E(X + Y) = E(X) + E(Y)$ $\qquad$ $V(X + Y) = V(X) + V(Y)$

Thus, students will need to develop an appreciation that while linearity is a desirable property for the purposes of computation and simplification, it is also a special one which applies in *some* contexts, but *not* in others.

FORMS INVOLVING $f(x + y)$, $f(x) + f(y)$ AND OTHER EXPRESSIONS

The follow discussion investigates whether different functions are, or are not, solutions to the Cauchy equation, and also whether there are other simple functional equations involving $f(x + y)$ or $f(x) + f(y)$, $f(x)$ and $f(y)$ that have particular solutions involving real differentiable functions. From the earlier work in Chapters 1 and 2, it follows that f: $R \to R$, $f(x) = x^2$, is *not* a solution to the Cauchy functional equation, while $h: R \to R$, $h(x) = 2x$ *is* a solution to this functional equation.

EXAMPLE 4.1

An approach to finding a possible solution to a functional equation is to start with the general form of a particular known type of function, and see what conditions (if any) are required for this to be the case. While this approach does *not* answer the question of identifying all possible solutions to a functional equation, it does identify those solutions of a particular kind. A similar approach is taken in various aspects of investigating the solution of certain types of differential equation.

For example, consider the general linear function $f: R \to R$, $f(x) = ax + b$ where the parameters a and b are real valued constants. *If* this form of function is a solution to the Cauchy functional equation, *then*:

$$\begin{array}{lrcl} & f(x+y) & = & f(x) + f(y) \\ \Rightarrow & a(x+y) + b & = & ax + b + ay + b \\ \Rightarrow & ax + ay + b & = & ax + ay + 2b \end{array}$$

This will only be the case when $b = 0$. Thus, any linear function of the form $f: R \to R$, $f(x) = ax$, where a is a real valued constant, will be a solution to the Cauchy equation, including the trivial case of the constant function with rule $f(x) = 0$, that is, the case where $a = b = 0$.

However, this is the only constant function, $f(x) = k$, where k is a real constant, that is a solution; since if k is non-zero, then by definition $f(x + y) = k$, but $f(x) + f(y) = k + k = 2k$. Thus, if, for example, $k = 3$, and hence $f: R \to R$, $f(x) = 3$, then it is the case by definition of the function that $f(x + y) = 3$ for all real combinations of x and y; however, $f(x) + f(y) = 3 + 3 = 6$, for all real combinations of x and y, as can be seen from Figure 4.1.

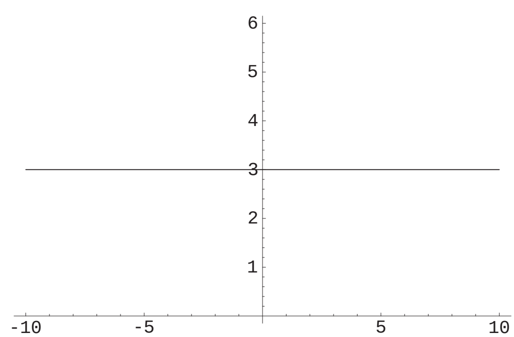


Figure 4.1: Graph of the constant function $f(x) = 3$

If a similar approach is tried with polynomials of higher degree, the corresponding set of simultaneous equations that specify the required conditions only yield a solution when each of the coefficients of powers of x, except the linear term, are zero. That is, in each case they reduce to the previous solutions.

EXAMPLE 4.2

Consider as an alternative, the set of power functions, $f: R \rightarrow R$, $f(x) = x^n$, where n is a positive integer. For the cases where $n = 1$ and $n = 2$, positive and negative results have already been established respectively. When $n > 2$, an algebraic approach can be used to show that $f(x + y) \neq f(x) + f(y)$. Table 4.1 provides the expansion of $f(x + y) = (x + y)^n$ for $n = 0$ to 6.

Table 4.1: Expansion of $f(x + y) = (x + y)^n$ for $n = 0$ to 6

0	1
1	$x+y$
2	$x^2+2xy+y^2$
3	$x^3+3x^2y+3xy^2+y^3$
4	$x^4+4x^3y+6x^2y^2+4xy^3+y^4$
5	$x^5+5x^4y+10x^3y^2+10x^2y^3+5xy^4+y^5$
6	$x^6+6x^5y+15x^4y^2+20x^3y^3+15x^2y^4+6xy^5+y^6$

From this it can be seen that for $n > 2$, there are 'middle terms' in the expansion of $(x + y)^n$, which, when x and y are non-zero, are themselves non-zero. Thus, for most combinations of values of x and y, $f(x + y)$ will not be equal to $f(x) + f(y)$. The general case for this can be argued from Pascal's triangle or the binomial expansion of $(x + y)^n$:

$$(x+y)^n = \binom{n}{n}x^n y^0 + \binom{n}{n-1}x^{n-1}y^1 + \binom{n}{n-2}x^{n-2}y^2 + \ldots$$
$$\binom{n}{2}x^2 y^{n-2} + \binom{n}{1}x^1 y^{n-1} + \binom{n}{0}x^0 y^n$$

Indeed, this form plays an important role in determining the rule for the derivative of $f(x) = x^n$ where $n \in N$; from the first principles definition of the derivative function:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

$$f'(x) = \lim_{h \to 0} \frac{f(x+h)^n - x^n}{h}$$

$$= \lim_{h \to 0} \frac{x^n + \binom{n}{n-1}x^{n-1}h + \binom{n}{n-2}x^{n-2}h^2 + \ldots + \binom{n}{0}h^n - x^n}{h}$$

$$= \lim_{h \to 0} \frac{\binom{n}{n-1}x^{n-1}h + \binom{n}{n-2}x^{n-2}h^2 + \ldots + \binom{n}{0}h^n}{h}$$

$$= \lim_{h \to 0} \left(nx^{n-1} + \binom{n}{n-2}x^{n-2}h + \ldots h^{n-1} \right)$$

In this case, the 'middle terms' of the binomial expansion in the numerator tend to zero as h tends to the limiting value of zero since each of these terms contains a non-zero power of h.

For some functions there are algebraically equivalent expressions for $f(x + y)$ in terms of a simple combination of $f(x)$ and $f(y)$, even if they are not of the form $f(x) + f(y)$.

EXAMPLE 4.3

In some cases it is possible to identify functions that are solutions to functional equations involving one of $f(x + y)$ or $f(x) + f(y)$. For example, consider the function $f: R \to R$, $f(x) = 2^x$.

If the values $x = 3$ and $y = 4$ are assigned, it is clear that $f(x + y) \neq f(x) + f(y)$ since $f(3 + 4) = f(7) = 2^7 = 128$ while $f(3) + f(4) = 2^3 + 2^4 = 8 + 16 = 24$. This can also be seen by inspection of the graph of f shown in Figure 4.2 by considering sample vertical and horizontal lines for coordinates:

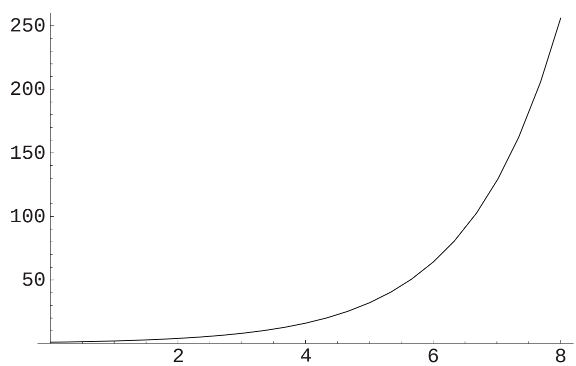


Figure 4.2: Graph of the exponential function with rule $f(x) = 2^x$

Hence, this function is *not* a solution to the Cauchy equation.

However, it is the case that $8 \times 16 = 2^3 \times 2^4 = 2^{3+4} = 2^7 = 128$, as this is an instance of the general law of exponents $a^x \times a^y = a^{x+y}$ for the assignment of values $a = 2$, $x = 3$ and $y = 4$.

The corresponding functional equation, for which *exponential functions*, $f: R \to R$, $f(x) = a^x$, where $a \in R^+$, provide a solution is $f(x + y) = f(x) \times f(y)$. That is, if $f: R \to R$, $f(x) = a^x$, where $a \in R^+$, then $f(x + y) = a^{x+y} = a^x \times a^y = f(x) \times f(y)$.

Exponential functions are therefore solutions to the functional equation $f(x + y) = f(x) \times f(y)$. The exponential function $f: R \to R$, $f(x) = e^x$ is *also* solution to the functional equation $f(x) = f'(x)$.

There is an *isomorphism*, which is Greek for same (*iso*) form (*morphism*), of mathematical structures between the *multiplicative group* of exponential terms to a given base, that is, the structure comprising $\{a^x, a \in (1, \infty)$ and $x \in R$ under the operation '$\times$'$\}$; and the *additive group* of logarithmic terms to the same base, that is, the structure comprising $\{\log_a(x), a \in (1, \infty)$ and $x \in R^+$ under the operation '$+$'$\}$.

Key properties of this isomorphism are summarised in Table 4.1.

Table 4.1: Summary of isomorphism between exponential and logarithmic structures

Exponential structure: $f(x) = a^x$	*Logarithmic structure: $g(x) = \log_a(x)$*
Natural domain = R, range = R^+	Natural domain = R^+, range = R
$y = a^x \Leftrightarrow x = \log_a(y)$	$y = \log_a(x) \Leftrightarrow x = a^y$
$a^0 = 1 \Leftrightarrow f(0) = 1$	$\log_a(1) = 0 \Leftrightarrow g(1) = 0$
$f(x + y) = f(x) \times f(y)$	$f(x \times y) = f(x) + f(y)$
$f(g(x)) = x$	$g(f(x)) = x$

The last row is a statement of the functional equation that defines a pair of *inverse functions*, the graphs of these two function exhibit reflectional symmetry in the line $y = x$, as shown in Figure 4.3 for the particular case where $a = 2$.

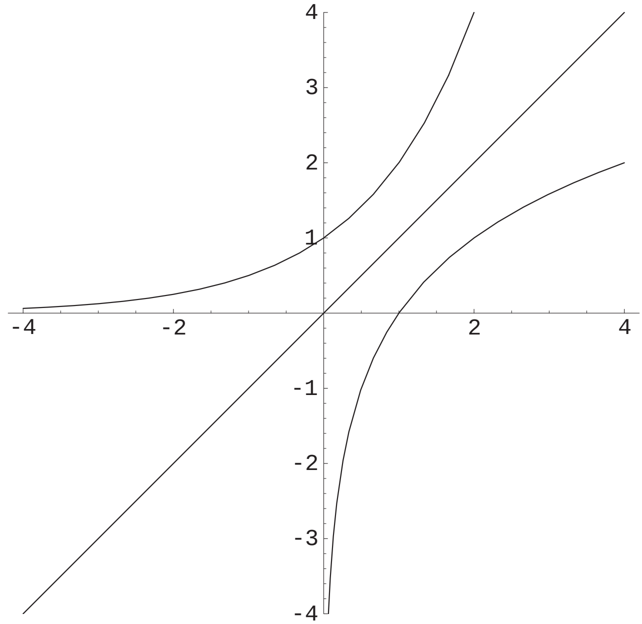


Figure 4.3: Graph of the inverse exponential-logarithmic pair of functions, base $a = 2$

Historically, this relationship was used to carry out computations before efficient mechanical and electronic calculators became widely available. The steps in the calculation require a mapping from the multiplicative structure to the additive structure via the 1–1 mapping of the logarithmic function (*taking logarithms*) carrying out addition (*of exponents = logarithms*), then applying the inverse mapping of exponentiation (*taking anti-logarithms*). This is illustrated in Figure 4.4, working in base 10.

27	×	**534**	=	**14418**
↓ (log function)		↓ (log function)		↑ (exponential function)
$10^{1.431363764}$	×	$10^{2.727541257}$	=	$10^{4.158905021}$

Figure 4.4: Example of multiplication using logarithms

The 'addition' occurs in the second last step where:

$$\log_{10}(27) + \log_{10}(534) = 1.431363764 + 2.727541257$$
$$= 4.158905021 = \log_{10}(27 \times 534)$$

FUNCTIONAL EQUATIONS AND ALGEBRAIC EQUIVALENCE

Various 'rules of algebra' that students encounter in their mathematical studies, can be considered in terms of relationships between $f(x + y)$, $f(x - y)$, $f(x \times y)$ or $f\left(\frac{x}{y}\right)$ and $f(x)$ and $f(y)$ for well-known functions such as power functions, exponential functions, logarithmic functions and circular functions. Each of these functions can be related to a set of functional equations for which they are a solution. The preceding example linked exponential and logarithm functions to particular functional equations that they satisfy: the exponent law $a^{x+y} = a^x \times a^y$ and the logarithm law $\log_a(xy) = \log_a(x) + \log_a(y)$.The following examples look at several other sorts of functional equations that can be used to describe algebraic equivalence with respect to various functions.

EXAMPLE 4.4

Consider the 'surd rules of algebra' for non-negative real numbers x and y:

$\sqrt{xy} = \sqrt{x}\sqrt{y}$ and $\sqrt{\frac{x}{y}} = \frac{\sqrt{x}}{\sqrt{y}}$ where $y > 0$

$\sqrt{x+y} \neq \sqrt{x} + \sqrt{y}$ and $\sqrt{x-y} \neq \sqrt{x} - \sqrt{y}$

as indicated by the counter-examples:

$\sqrt{16+9} = 5$ but $\sqrt{16} + \sqrt{9} = 4 + 3 = 7$

and

$\sqrt{16-9} = \sqrt{7} \approx 2.646$ but $\sqrt{16} - \sqrt{9} = 4 - 3 = 1$

In terms of the function f: $[0, \infty) \to R$, where $f(x) = \sqrt{x}$, the first two of these correspond to this function being a solution of the functional equations:

$f(x \times y) = f(x) \times f(y)$ and $f\left(\frac{x}{y}\right) = \frac{f(x)}{f(y)}$

but *not* a solution of the functional equations:

$f(x + y) = f(x) + f(y)$ and $f(x - y) = f(x) - f(y)$.

However, it should also be noted that $g(x) = \frac{1}{x}$ where x is a non-zero real number, and $h(x) = | x |$ are also solutions (or not) to the same set of functional equations, with suitable restriction to ensure there is a non-zero divisor in the second functional equation. Indeed, if x is a positive real number, then $f(x) = x^q$, where q is a non-zero rational number, is a solution to the first two functional equations, but not the second two functional equations.

If different combinations of functional terms are 'mixed and matched' there are also the following solutions to particular functional equations that characterise the other 'exponent and logarithm' laws:

$f(x - y) = \dfrac{f(x)}{f(y)}$ the exponent law $a^{x-y} = \dfrac{a^x}{a^y}$

$f\left(\dfrac{x}{y}\right) = f(x) - f(y)$ the logarithm law $\log_a\left(\dfrac{x}{y}\right) = \log_a(x) - \log_a(y)$

EXAMPLE 4.5

For the basic circular functions $\sin(x)$, $\cos(x)$ and $\tan(x)$, it is certainly the case that these functions are *not* solutions to the Cauchy equation. For example, in the case of the sine function, using the well-known exact value $\sin\left(\dfrac{\pi}{2}\right)$, it is clear that $\sin(x + y) \neq \sin(x) + \sin(y)$ since

$$\sin\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = \sin(\pi) = 0 \text{ but } \sin\left(\frac{\pi}{2}\right) + \sin\left(\frac{\pi}{2}\right) = 1 + 1 = 2.$$

Similarly, counter-examples using exact values can be found for the cosine and tangent functions with respect to this and other functional equations such as $f(x \times y) = f(x) \times f(y)$ and $f\left(\dfrac{x}{y}\right) = \dfrac{f(x)}{f(y)}$.

For example, $\cos(4\pi) \neq \cos(4)\cos(\pi)$ and $\tan\left(\dfrac{\pi}{3}\right) \neq \dfrac{\tan(\pi)}{\tan(3)}$. This can be readily determined from a knowledge of exact values for these circular functions, or inspection of the corresponding graphs.

There are, however, well-known algebraic equivalences (usually called *trigonometric identities*) for these circular functions. These can be obtained in several ways, and in the following discussion matrices are used to simultaneously obtain expressions for $f(x + y)$ when f is the sine or cosine function. These can be subsequently used to obtain the relevant functional equation for $\tan(x + y)$ from the definition:

$$\tan(x + y) = \frac{\sin(x + y)}{\cos(x + y)}$$

If the images of the points (1, 0) and (0, 1) under a given matrix transformation M of the cartesian plane are known, then the matrix is completely determined. This is because these two points form the basis of the vector space for all coordinate vectors in the plane as $(x, y) = x(1, 0) + y(0, 1)$ under scalar multiplication and vector addition as shown, for example for (5, 3) in Figure 4.5.

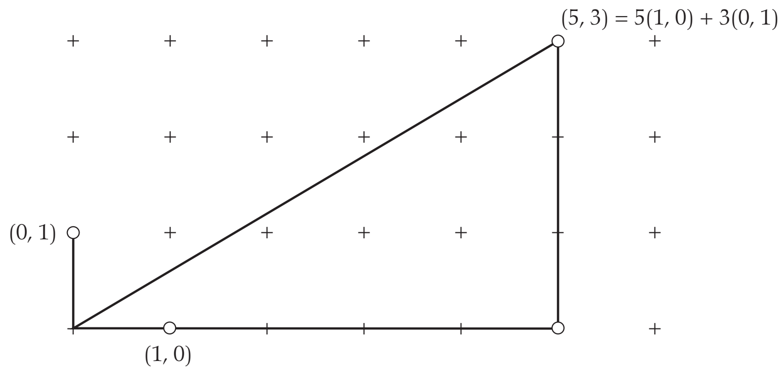


Figure 4.5: Diagram of coordinate vector in the plane as (5, 3) = 5(1, 0) + 3(0, 1)

Using matrix notation and column vectors, if the images of (1, 0) and (0,1) are respectively (a, c) and (b, d) then:

$$M\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}a\\c\end{bmatrix} \quad \text{and} \quad M\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}b\\d\end{bmatrix}$$

Combining these two matrix equations into a single matrix equation gives:

$$M\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}a & b\\c & d\end{bmatrix}$$

Consider the following diagram where the unit line segments from the origin with endpoints at (1, 0) and (0, 1) respectively are both rotated through an angle θ anticlockwise, as shown in Figure 4.6.

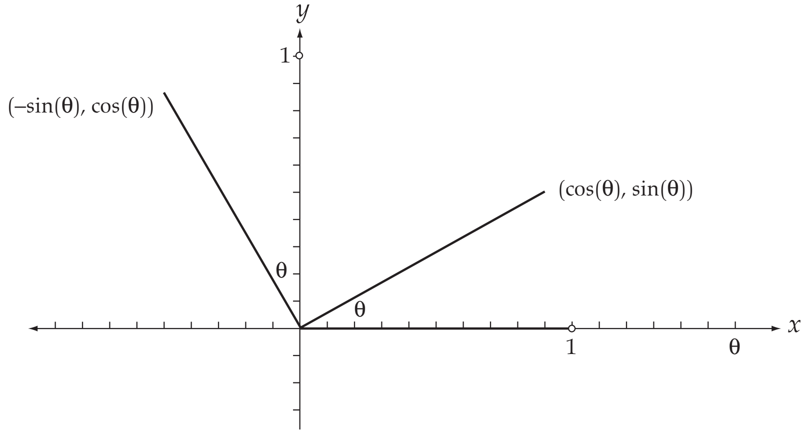


Figure 4.6: Rotation of (1, 0) and (0, 1) through angle of θ anticlockwise

If a rotation anticlockwise through an angle θ about the origin (0, 0) is designated by R_θ, then Figure 4.6 and basic trigonometry of right-angled triangles shows that:

$$R_\theta \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \quad \text{and} \quad R_\theta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$$

In matrix form this is equivalent to:

$$R_\theta \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

How are the compound angle formulas for sine *and* for cosine obtained from this? Simply by observing that rotation about the origin through an angle of θ + φ can be obtained in two equivalent ways—a single rotation through θ + φ *or* a rotation through θ followed by a rotation through φ.

The former is given by the matrix for $R_{\theta+\varphi}$ while the latter is given by the product $R_\varphi R_\theta$. Thus, in matrix form:

$$R_{\theta+\varphi} = \begin{bmatrix} \cos(\theta+\varphi) & -\sin(\theta+\varphi) \\ \sin(\theta+\varphi) & \cos(\theta+\varphi) \end{bmatrix}$$

while

$$R_\varphi R_\theta = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\varphi)\cos(\theta) - \sin(\varphi)\sin(\theta) & \cos(\theta) \times -\sin(\theta) - \sin(\varphi)\cos(\theta) \\ \sin(\varphi)\cos(\theta) + \cos(\varphi)\sin(\theta) & -\sin(\varphi)\sin(\theta) + \cos(\varphi)\cos(\theta) \end{bmatrix}$$

So, by the equality of matrices, which requires corresponding elements in the arrays to be the same:

$$\sin(\theta + \varphi) = \sin(\theta)\cos(\varphi) + \cos(\theta)\sin(\varphi)$$

and

$$\cos(\theta + \varphi) = \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi)$$

If the identity $\cos(\theta) = \sin(\theta + \frac{\pi}{2})$ is also used, then the identity for $\sin(\theta + \varphi)$ can be expressed using the functional equation:

$$f(x + y) = f(x)\cos(y) + \cos(x) f(y) = f(x) f(y + \tfrac{\pi}{2}) + f(x + \tfrac{\pi}{2}) f(y)$$

Similar expressions can be obtained for cosine and tangent using symmetry properties of these circular functions.

SUMMARY

- The functional equation $f(x + y) = f(x) + f(y)$ is called the *Cauchy equation* and has the solution $f: R \to R$, $f(x) = ax$, where a is a real valued constant, for real differentiable functions of a single variable.
- Most real continuous functions are *not* solutions of the functional equations $f(x + y) = f(x) + f(y)$ and $f(x - y) = f(x) - f(y)$, the assumption of this linearity condition is a common error in algebraic manipulations.
- $(x + y)^n \neq x^n + y^n$ for $n \in N$ and $n > 1$
- $a^{x+y} \neq a^x + a^y$
- $\log_a(x + y) \neq \log_a(x) + \log_a(y)$
- $\sin(x + y) \neq \sin(x) + \sin(y)$ and similarly for the cosine and tangent functions
- The linearity property does apply in several contexts, such as differentiation and integration of real functions, linear (affine) transformations of the cartesian plane, and expectation and variance of the sum of two independent random variables.

- Functions of the form $f: R \to R$, $f(x) = x^q$ (power functions, $q \neq 1$ and $q \in Q$), or the form $f: R \to R$, $f(x) = a^x$, where $a \in R^+$ (exponential functions) are **not** solutions of the *Cauchy equation*.
- Power function of the form $f: R^+ \to R$, $f(x) = x^q$ where $q \in Q$ are solutions of the functional equations $f(x \times y) = f(x) \times f(y)$ and $f\left(\frac{x}{y}\right) = \frac{f(x)}{f(y)}$.
- Exponential functions of the form $f: R \to R$, $f(x) = a^x$, where $a \in R^+$ are solutions of the functional equations $f(x + y) = f(x) \times f(y)$ and $f(x - y) = f\left(\frac{x}{y}\right)$.
- Logarithmic functions of the form $f: R^+ \to R$, $f(x) = \log_a(x)$ where $a \in R^+$ are solutions of the functional equations $f(x \times y) = f(x) + f(y)$ and $f\left(\frac{x}{y}\right) = f(x) - f(y)$.
- Circular functions satisfy functional equations corresponding to trigonometric identities (algebraic equivalences) for $\sin(x + y)$, $\cos(x + y)$, $\tan(x + y)$ and related forms.

STUDENT ACTIVITY 4.1

Use tables, graphs or algebraic reasoning to decide whether or not each of the following functions is a solution to the functional equation $f(x + y) = f(x) + f(y)$.

a $f(x) = x$
b $f(x) = 2x^2$
c $f(x) = x^3$
d $f(x) = x^4$
e $f(x) = 3$
f $f(x) = \sqrt{x}$
g $f(x) = \sin(x)$
h $f(x) = \cos(x)$
i $f(x) = \frac{1}{x}$

STUDENT ACTIVITY 4.1 (CONTINUED)

j $f(x) = \frac{1}{x^2}$
k $f(x) = e^x$
m $f(x) = \log_{10}(x)$
n $f(x) = |x|$

STUDENT ACTIVITY 4.2

a Show that the function $f: R \longrightarrow R$, $f(x) = ax$ is a solution of the functional equation $f(x - y) = f(x) - f(y)$, and provide a graphical interpretation of this result.
b Compare this with the evaluation and graphical interpretation of $f(x) - f(y)$ for $f: R \longrightarrow R$, $f(x) = ax + b$. (Hint: think of x as x_1 and y as x_2 in the usual calculation of gradient of a non-vertical straight line from the coordinates of two distinct points.)

STUDENT ACTIVITY 4.3

a Let $g: R\backslash\{0\} \longrightarrow R$, $g(x) = \frac{1}{x}$. Show that there are *no* integer values of x and y for which $g(x + y) = g(x) + g(y)$.
b Find a suitable expression for $g(x + y)$ in terms of $g(x)$ and $g(y)$.

STUDENT ACTIVITY 4.4

Identify differentiable real functions of a single variable that are solutions to the following functional equations:

a $f\left(\frac{x+y}{2}\right) = \frac{f(x) + f(y)}{2}$
b $f(x + y) + f(x - y) = 2f(x) \times f(y)$
c $f(x + y) \times f(x - y) = f(x)^2 - f(y)^2$

Find a functional equation for which the following are a solution:

d $f(x) = \tan(x)$
e $f(x) = a \times e^{cx^2}$

STUDENT ACTIVITY 4.5

Show that the function $f: R^+ \to R$, $f(x) = x^q$, where q is a rational number, satisfies the functional equations:

$$f(x \times y) = f(x) \times f(y) \text{ and } f\left(\frac{x}{y}\right) = \frac{f(x)}{f(y)}$$

but *not* the functional equations:

$$f(x + y) = f(x) + f(y) \text{ and } f(x - y) = f(x) - f(y).$$

STUDENT ACTIVITY 4.6

Investigate the historical development of natural logarithms, and related functional equations. How is the logarithm function *defined* in modern mathematics?

STUDENT ACTIVITY 4.7

a Let X and Y be 2×2 matrices that have inverses under matrix multiplication. If $X^2 = X \times X$, show that, in general $(X \times Y)^2 \neq X^2 \times Y^2$ but that $(X \times Y)^{-1} = Y^{-1} \times X^{-1}$. For what 2×2 matrices does $(X \times Y)^2$ equal $X^2 \times Y^2$?

b Show that for any linear transformation of the plane, a pair of parallel lines will be mapped onto an image pair of parallel lines.

STUDENT ACTIVITY 4.8

The structure for differentiation of functions of a real variable is based on the properties of the derivative that for two differentiable real functions f and g:

- $\boldsymbol{D}[f(x) + g(x)] = \boldsymbol{D}[f(x)] + \boldsymbol{D}[g(x)]$ and
- $\boldsymbol{D}[f(x) \times g(x)] = f(x) \times \boldsymbol{D}[g(x)] + g(x) \times \boldsymbol{D}[f(x)]$

Explain why it is *not* true that $\boldsymbol{D}[f(x) \times g(x)] = \boldsymbol{D}[f(x)] \times \boldsymbol{D}[g(x)]$ or $\boldsymbol{D}\left[\frac{f(x)}{g(x)}\right] = \frac{\boldsymbol{D}[f(x)]}{\boldsymbol{D}[g(x)]}$.

STUDENT ACTIVITY 4.9

Use the equivalences:

$$\sin(\theta + \varphi) = \sin(\theta)\cos(\varphi) + \cos(\theta)\sin(\varphi)$$

and

$$\cos(\theta + \varphi) = \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi)$$

STUDENT ACTIVITY 4.9 (CONTINUED)

to obtain equivalences for:

a $\tan(\theta + \varphi)$
b $\sin(2\theta)$, $\cos(2\theta)$ and $\tan(2\theta)$
c $\sin(\theta - \varphi)$, $\cos(\theta - \varphi)$ and $\tan(\theta - \varphi)$

Represent each of these identities using a suitable functional equation.

References

Aigner, M & Zeigler, EM 2004, *Proofs from the book* (3rd edn), Springer-Verlag, New York.

Binmore, KG 1977, *Mathematical analysis: A straightforward approach*, Cambridge University Press, Melbourne.

Lipschutz, S 1974, *Linear algebra*, McGraw-Hill, New York.

Martin, GE 1982, *Transformation geometry: An introduction to symmetry*, Springer-Verlag, New York.

Websites

http://eqworld.ipmnet.ru/ – Russian Academy of Sciences
A free website on *The World of Mathematical Equations*. Provides a rigorous approach to solution methods. See in particular 'Functional Equations' in the Exact Solutions, Methods and education sections, and the Functional Equations – Index.

http://www.maths.usyd.edu.au/u/richardc/traffic.html – School of Mathematics and Statistics, University of Sydney
A paper on traffic flow that involves functional equations and probability distributions.

http://www.maths.mq.edu.au/~alf/PaperfoldingTalk.pdf – Department of Mathematics, Macquarie University
A mathematics paper based on functional equations that arise in the context of paper folding.

CHAPTER 5

DIFFERENCE EQUATIONS

Since the 1980s, *difference equations*, or *recurrence relations*, have increasingly become part of the senior secondary mathematics curriculum where there is a *discrete mathematics* emphasis, often in conjunction with financial or business applications. Difference equations are a particular kind of functional equation where the independent variable takes natural number values, and where the dependent variable forms a countable sequence of values or *terms* as they are commonly called. Thus, the value of an investment over a set of fixed time intervals would typically be described by some sort of difference equation or recurrence relation. These are characterised by the property that a given value is computed from a previous value or values in the sequence.

This process typically requires the ability not only to carry out lots of computations efficiently, but also to remember the results of *previous* calculations for use in a *current* calculation, and so on. Hand-held technology such as graphics or CAS calculators, or computer-based technology such as spreadsheets and computer algebra systems, are well suited to the analysis and solution of difference equations or recurrence relations, and the representation of corresponding sequences by lists, tables and graphs. More broadly, the notion of *recursion* is closely related to the development of number theory, computable functions, and its application to computation in general (see Enderton 1979; Mendelson 1979; Cohen 2002). Indeed, some form of recursion or iteration is at the heart of most forms of practical computation carried out by technology.

More formally, functional equations that have solutions which are functions with the domain $N^+ = \{1, 2, 3, \dots\}$ are called *difference* equations or *recurrence* relations. If f is such a function, then the list of the range values of the function $\{f(1), f(2), f(3), f(4), \dots\}$ is called a *sequence*, and $f(n)$ is referred to as the nth term of the sequence. For difference or *recursion* equations, a given term is defined in relation to the previous term or terms, hence the name 'recursion', from the Latin *recurso*—to run back. Thus, a solution to a functional equation that is a difference equation is a function f, with natural domain, N^+, and which generates a sequence of values $\{ f(1), f(2), f(3), f(4), \dots \}$ that matches a given sequence. An extensive range of applications of sequences generated by solution functions to difference equations, in particular financial applications such as annuities, loan amortisation and saving with a fixed periodic investment, can be found in Hodgson and Leigh-Lancaster (1990).

Two well-known types of sequences that students typically meet in their senior secondary mathematics studies are *arithmetic* and *geometric* sequences. Indeed, their work on linear and exponential functions is, in many respects, a study of these functions on an integer subset of their real number maximal domain, with continuity implied by the use of graphical representation using lines and smooth curves. *Arithmetic* and *geometric* sequences are special cases of *first order linear* difference equations, that is, a given term is described in relation to its previous term only and that relation is linear.

Several second order difference equations that students may also come across are the difference equations that lead to quadratic polynomial functions (through difference tables with a constant second difference, a linear second order difference equation), *Fibonacci* sequences (also a linear second order difference equation) where the *logistic* equation is a non-linear first order difference equation.

ARITHMETIC SEQUENCES

The sequence $S = \{2, 6, 10, 14, \dots\}$ is an *arithmetic* sequence, specified completely by the solution to the functional equation $f(x + 1) = f(x) + 4$, over domain N^+, that satisfies the initial condition $f(1) = 2$. The value $f(x + 1) - f(x) = 4$ is called the *common difference* of the sequence, while $f(1) = 2$ is said to be the *first term*. When the natural domain is known to be N^+, the independent variable is usually designated by n rather than x.
The sequence is generated as follows:

$$\begin{aligned} f(1) &= && && 2 \\ f(2) &= f(1) + 4 &&= 2 + 4 &&= 6 \\ f(3) &= f(2) + 4 &&= 6 + 4 &&= 10 \end{aligned}$$

$$f(4) = f(3) + 4 = 10 + 4 = 14$$

$$\begin{array}{l} . \\ . \\ . \end{array}$$

and so on.

Now in this case it is not too difficult to see that an explicit formula for $f(n)$ can be determined, since this is effectively a linear function with domain N. Indeed, $f(n) = 2 + (n - 1) \times 4 = 4n - 2$, where $n = 1, 2, 3, \ldots$

More generally, if the initial condition is $f(1) = a$ and $f(n + 1) - f(n) = d$, then the explicit formula for $f(n)$ will be $f(n) = a + (n - 1) \times d = dn - d + a$, where $n = 1, 2 \ldots$ The corresponding sequence then becomes $\{a, a + d, a + 2d, a + 3d, a + 4d, \ldots\}$.

The same type of sequence can also be generated by repeated composition, or *nesting* of a function f, starting from the initial evaluation of a given, or *fixed*, point. Thus, starting with a fixed point $x = a$, the sequence $\{a, f(a), f(f(a)), f(f(f(a))), f(f(f(f(a)))) \ldots\}$ is generated. This process is also called *fixed point iteration*. Some authors prefer to start with $n = 0$, others with $n = 1$, depending on whether they wish to emphasise the first value of n corresponding to the first term in the sequence or the first application of the recursion process.

The arithmetic sequence $S = \{2, 6, 10, 14, \ldots\}$ is generated by fixed point iteration as follows: let $a = 2$, and $f(n) = n + 4$. Then successive iterations give:

$$\begin{array}{llll} a & = & & 2 \\ f(a) & = 2 + 4 & = & 6 \\ f(f(a)) & = 2 + 4 + 4 & = & 10 \\ f(f(f(a))) & = 2 + 4 + 4 + 4 & = & 14 \\ . & & & . \\ . & & & . \\ . & & & . \end{array}$$

and so on, as previously.

CAS such as *Mathematica* are well suited to this type of computation, and have corresponding functionality built in, in this case in the form of the functions **Nest** and **NestList**. For an arbitrary function, f, and fixed point, a, nesting 5 times gives:

```
Nest[f,a,5]
f[f[f[f[f[a]]]]]
```

while nest-listing gives the sequence of all the previous fixed point iterations as well:

```
NestList[f,a,5]
{a,f[a],f[f[a]],f[f[f[a]]],f[f[f[f[a]]]],f[f[f[f[f[a]]]]]}
```

for the sequence S this has the particular form:

```
f[n_] := n + 4
{2,6,10,14,18,22}
```

which can then be graphed as a plot of this list of values as shown in Figure 5.1:

```
ListPlot[NestList[f,2,5],PlotRange → {0,30}] :
```

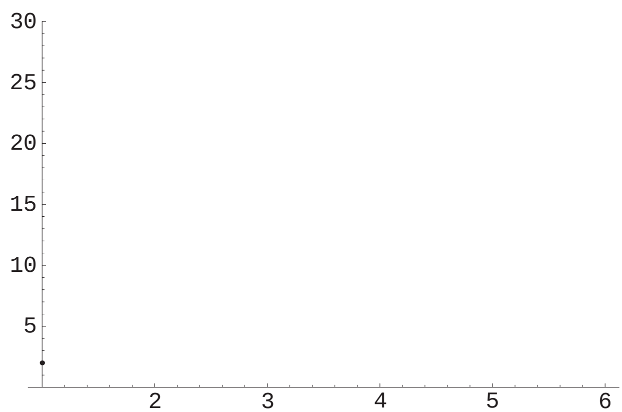


Figure 5.1: Plot of the first six values of sequence *S*

The graph of this function is clearly linear, as would be expected for an arithmetic sequence. If the function is left arbitrary, it is not possible to produce a corresponding graph. The general arithmetic sequence can be obtained by defining $f(n) = n + d$, and using $f(1) = a$ as the fixed point:

```
f[n _] := n + d
{a,a + d,a + 2d,a + 3d,a + 4d,a + 5d}
```

The *constant* sequence $\{c, c, c \ldots\}$, where $c \in R$, is a special case of an arithmetic sequence where $f(0) = c$ and $d = 0$.

GEOMETRIC SEQUENCES

The sequence $G = \{2, 6, 18, 54, \ldots\}$ is a *geometric* sequence, specified completely by the solution to the functional equation $f(x + 1) = 3 \times f(x)$, over domain N^+, and which satisfies the initial condition $f(1) = 2$. Again, using n to designate the independent variable of a sequence, the value $\dfrac{f(n+1)}{f(n)} = 3$ is called the *common ratio* of the sequence, while $f(1) = 2$ is said to be the *first term*. The sequence is generated as follows:

$$\begin{array}{llllllll}
f(1) & = & & & & & 2 \\
f(2) & = & f(1) \times 3 & = & 2 \times 3 & = & 6 \\
f(3) & = & f(2) \times 3 & = & 6 \times 3 & = & 18 \\
f(4) & = & f(3) \times 3 & = & 10 \times 3 & = & 54 \\
. & & & & & & . \\
. & & & & & & . \\
. & & & & & & .
\end{array}$$

and so on.

Thus, while both the previous *arithmetic* sequence and this *geometric* sequence have the same first two terms, all the other terms will be different. They are, in general, different types of sequences. In this case it is not too difficult to see that an explicit formula for $f(n)$ can be determined, since this is effectively an exponential function with domain N^+.

Indeed, $f(n) = 2 \times 3^{n-1} = \frac{2}{3} \times 3^n$, where $n = 1, 2, 3 \ldots$

More generally, if the initial condition is $f(1) = a$ and $\dfrac{f(n+1)}{f(n)} = r$ then the explicit formula for $f(n)$ will be $f(n) = a \times r^{n-1} = \dfrac{a}{r} \times r^n$, where $n = 1, 2, 3 \ldots$

The corresponding sequence then becomes $\{ a, ar, ar^2, ar^3 \ldots \}$.

The same type of sequence can similarly also be generated by repeated composition, or *nesting* of a function f, staring from the initial evaluation of a given, or *fixed* point. Thus, starting with a fixed point $x = a$, the sequence $\{a, f(a), f(f(a)), f(f(f(a))), f(f(f(f(a)))) \ldots\}$ is generated, except that the form of the rule for f is different for a geometric sequence. This process is again called *fixed point iteration*.

For example, the geometric sequence $G = \{2, 6, 18, 54, \ldots\}$ is generated by fixed point iteration as follows. Let $a = 2$, and $f(n) = 3n$. Then successive iterations give:

$$\begin{array}{llll}
a & = & & 2 \\
f(a) & = & 2 \times 3 = & 6 \\
f(f(a)) & = & 2 \times 3 \times 3 = & 18 \\
f(f(f(a))) & = & 2 \times 3 \times 3 \times 3 = & 54 \\
. & & & . \\
. & & & . \\
. & & & .
\end{array}$$

and so on, as previously. Using the CAS *Mathematica*, the sequence is produced by listing the iterates:

```
NestList[f,2,5]
{2,6,18,54,162,486}
```

and can be graphed as shown in Figure 5.2.

```
ListPlot[NestList[f,2,5],PlotRange → {0,500}]:
```

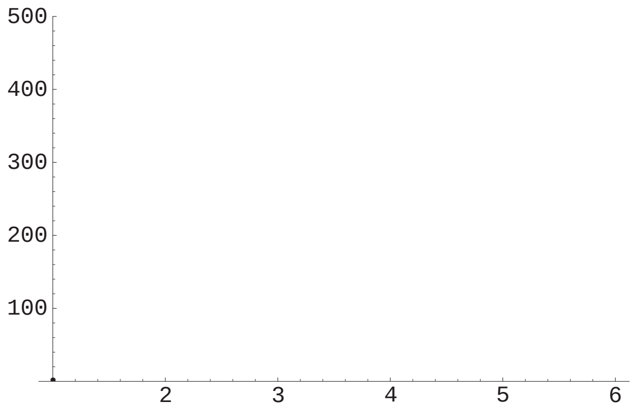


Figure 5.2: Plot of the first six values of sequence *G*

Again, if the function is left arbitrary, it is not possible to produce a corresponding graph; however, the general list of iterates for an arbitrary function is given. The general geometric sequence can be obtained by defining $f(n) = rn$, and using $f(1) = a$ as the fixed point:

```
f[n _] := rn
NestList[f,a,5]
{a,ar,ar^2,ar^3,ar^4,ar^5}
```

The behaviour of the sequence can be described in terms of the value of r, as summarised in Table 5.1.

Table 5.1: Behaviour of geometric sequence in terms of the value of *r*

Value of r	*Behaviour of sequence*
$r < -1$	oscillating and diverging
$r = -1$	oscillating between two constant values: $f(1) = a$ and $-f(1) = -a$
$-1 < r < 0$	oscillating and converging to zero
$r = 0$	first term $f(1) = a$ then a constant sequence of zero
$0 < r < 1$	exponential decay converging to zero
$r = 1$	constant sequence of $f(1) = a$
$r > 1$	diverging, exponential growth

Arithmetic sequences and geometric sequences are subsets of the corresponding continuous linear and exponential functions respectively. Thus, they will also be solutions to the functional equations that their continuous analogues satisfy.

FIRST ORDER LINEAR DIFFERENCE EQUATIONS

A first order linear difference equation is a functional equation of the form:

$$f(x + 1) = af(x) + b \quad \text{where } x \in N^+ \text{ and } a, b \in R$$

For example, consider the first order linear difference equation $f(n + 1) = 0.5\, f(n) + 1$, with first term $f(1) = 3$. The corresponding sequence of values can also be generated by repeated composition, or *nesting* of a function f starting from the initial value of a given, or *fixed* point, in this case $f(1) = 3$ to obtain $\{a, f(a), f(f(a)), f(f(f(a))), f(f(f(f(a)))) \ldots\}$.

The following computation using *Mathematica* evaluates this sequence for several iterations:

```
f[n_] := 0.5n + 1
NestList[f,3,10]
{3,2.5,2.25,2.125,2.0625,2.03125,2.01563,
2.00781,2.00391,2.00195,2.00098}
```

The graph of this function is shown in Figure 5.3.

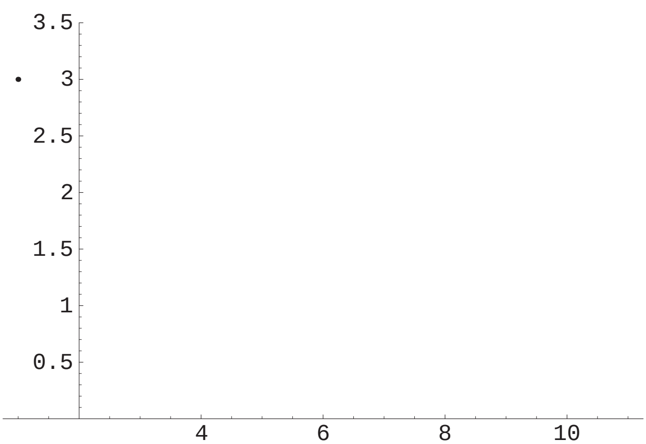


Figure 5.3: Plot of the first eleven values of $f(n + 1) = 0.5\, f(n) + 1$ where $f(1) = 3$

The explicit solution to the first order linear difference equation $f(n+1) = 0.5\,f(n) + 1$, $f(1) = 3$ can be obtained using CAS:

```
RSolve[{f[n] == 1/2 f[n − 1] + 1, f[1] == 3}, f[n], n]
{{f[n] → 2If[n ≥ 1,-1,-2^-n,0]}}
```

The corresponding general case solution is:

```
RSolve[{f[n] == af[n − 1] + b, f[1] == k}, f[n], n]
```

$$\left\{\left\{f[n] \to \text{If}\left[n \geq 1, \frac{-ab + a^n(b-k) + a^{1+n}k}{(-1+a)a}, 0\right]\right\}\right\}$$

This result can be derived by series analysis, using some algebraic manipulation and the sum of a geometric series. Assume $f(1)$ is given and $f(n+1) = af(n) + b$, then:

$f(2) = af(1) + b$

$f(3) = af(2) + b = a(\,af(1) + b) + b = a^2f(1) + ab + b$

$f(4) = af(3) + b = a(\,a^2f(1) + ab + b) + b = a^3f(1) + a^2b + ab + b$

.

.

.

$f(n) = a^{n-1}f(1) + (b + ab + a^2b + \ldots + a^{n-2}b)$

The term in brackets is a geometric sequence with first term, b, common ratio, a, and $n-1$ terms. Using the formula for the sum of the first n terms of a geometric sequence this becomes:

$$\frac{b(a^{n-1}-1)}{a-1}, \text{ where } a \neq 1$$

Thus, the general solution is $f(n) = a^{n-1}f(1) + \dfrac{b(a^{n-1}-1)}{a-1}$, where $a \neq 1$.

The general first order linear difference equation $f(n+1) = af(n) + b$, where $n \in N^+$ and $a, b \in R$ incorporates the cases of:

- the constant sequence $\{\,f(1), b, b, b \ldots\}$ where $a = 0$ and $f(n) = b$: for example, if $f(1) = -3$ and $f(n+1) = 0 \times f(n) + 7$, then the sequence generated is $\{-3, 7, 7, 7 \ldots\}$
- the arithmetic sequence $\{\,f(1), f(1) + d, f(1) + 2d, f(1) + 3d \ldots\}$ where $a = 1$ and $b = d$: for example, if $f(1) = -3$ and $f(n+1) = 1 \times f(n) + 7$, then the sequence generated is $\{-3, 4, 11 \ldots\}$

- the geometric sequence $\{f(1), f(1) \times r, f(1) \times r^2, f(1) \times r^3 \ldots\}$ where $a = r$ and $b = 0$: for example, if $f(1) = -3$ and $f(n + 1) = 7 \times f(n)$, then the sequence generated is $\{-3, -21, -147 \ldots\}$

It should be noted that the a used here is not the 'a' often used to designate the first term of an arithmetic or geometric sequence (this role is played by $f(1)$ in the current discussion) but the coefficient of $f(n)$ in the general first order linear difference equation $f(n + 1) = af(n) + b$.

If $f(1) = -3$ and the first order linear difference equation is $f(n + 1) = 7 \times f(n) + 7$, then the sequence generated is $\{-3, -14, -98 \ldots\}$ and is graphed in Figure 5.4.

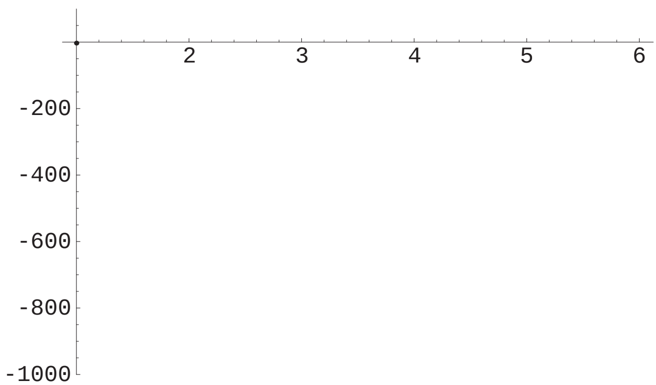


Figure 5.4: Plot of the first four values of $f(n + 1) = 7f(n) + 7$ where $f(1) = -3$

The corresponding explicit rule for this function is:

```
RSolve[{f[n] == 7f[n - 1] + 7, f[1] == -3}, f[n], n]
```

$$\left\{\left\{f[n] \to \text{If}\left[n \geq 1, \frac{1}{42}(-49 - 11\,7^{n}), 0\right]\right\}\right\}$$

DIFFERENCE TABLES FOR LINEAR AND QUADRATIC POLYNOMIAL FUNCTIONS

The linear function $f: N \to R, f(x) = ax + b$ is an explicit solution to the difference equation where $f(x + 1) = f(x) + a$ and $f(0) = b$. The rule of the explicit solution can be readily obtained using a finite difference table, which consists of a sequence of several consecutive values of the function for which a corresponding sequence of differences between consecutive values is obtained.

Consider a table of values for the function f where $x = 0, 1, 2, 3, 4$ and the corresponding differences between consecutive values (that is, the list of *first* differences) as shown in Table 5.2.

Table 5.2: First differences for a linear function

x	*f(x)*	*First difference*
0	b	
		a
1	$a + b$	
		a
2	$2a + b$	
		a
3	$3a + b$	
		a
4	$4a + b$	

The *constant* first difference implies the function is linear, and its graph would have gradient a with y-axis intercept at $f(0) = b$. The use of such a table is well known, for example given the following set of value for x and $f(x)$, and calculating the first difference gives:

x	*f(x)*	*First difference*
0	14	
		–3
1	11	
		–3
2	8	
		–3
3	5	
		–3
4	2	

from which it is readily deduced that the corresponding function would have the rule $f(x) = -3x + 14$.

For a quadratic function $f: N \rightarrow R, f(x) = ax^2 + bx + c$ the finite difference table has a non-constant first difference, but a constant second difference (that is, difference of first differences) as shown in Table 5.3.

Table 5.3: First and second differences for a quadratic function

x	f(x)	First difference	Second difference
0	c		
		$a + b$	
1	$a + b + c$		$2a$
		$3a + b$	
2	$4a + 2b + c$		$2a$
		$5a + b$	
3	$9a + 3b + c$		$2a$
		$7a + b$	
4	$16a + 4b + c$		

Consider any sub-sequence of three consecutive terms $\{f(n), f(n+1), f(n+2)\}$ of any sequence, then the two consecutive first differences formed from these are $f(n+2) - f(n+1)$ and $f(n+1) - f(n)$. The corresponding second difference is $f(n+2) - f(n+1) - (f(n+1) - f(n))$ which is equal to $f(n+2) - 2f(n+1) + f(n)$. For a quadratic function, the second difference is constant and equal to $2a$ that is:

$$f(n+2) - 2f(n+1) + f(n) = 2a \quad \text{or} \quad f(n+2) = 2f(n+1) - f(n) + k$$

for some non-zero real constant k. Thus, the quadratic function is a solution to this second order linear difference equation.

FIBONACCI SEQUENCES

An important class of sequences that satisfy the functional equation $f(x+2) = f(x+1) + f(x)$ are the Fibonacci and Lucas sequences, for example, $\{2, 4, 6, 10, 16 \ldots\}$ is a Lucas sequence. This is called a *second order* difference equation, since the *next* term of the sequence is defined in terms of the previous *two* terms. An alternative form of the functional equation is $f(x) = f(x-1) + f(x-2)$.

Historically, the Fibonacci sequence $\{1, 1, 2, 3, 5, 8 \ldots\}$ was studied by Leonardo of Pisa, also called Fibonacci (meaning 'son of Bonaccio'), in his book *Liber Abaci* (1202). Fibonacci was a merchant who had travelled extensively in the Orient, and used this sequence to model a population growth problem—that of pairs of breeding rabbits. The problem can be formulated as follows:

> Starting with a single pair of fertile rabbits, how many pairs of rabbits will be produced in a year if each pair gives birth to a new pair of rabbits each month, a new pair become fertile from the second month, fertile pairs are always productive and there are no deaths?

The solution to this problem can be obtained by using a tree diagram, labelling pairs of rabbits as *mature* (and hence breeding) or *new-born* (and hence fertile the following month) in each month and recording the results in a table as shown in Table 5.4. This is a good activity for students, who are quite likely to discover the Fibonacci sequence through this process anyway.

Table 5.4: The Fibonacci sequence (assuming births of new born pairs on the first day of the month)

Month	*Mature pairs*	*New-born pairs*	*Total pairs*
1	1	0	1
2	1	1	2
3	2	1	3
4	3	2	5
5	5	3	8
6	8	5	13
7	13	8	21
8	21	13	34
9	34	21	55
10	55	34	89
11	89	55	144
12	144	89	233
13	233	144	377

To see why this sequence satisfies the functional equation $f(n + 2) = f(n + 1) + f(n)$, assume that in a given month, n, there are a pairs of rabbits (combined mature and new-born), and in the following month, $n + 1$, there are b pairs of rabbits (combined mature and new-born). In the next month, $n + 2$, there will be $a + b$ pairs of rabbits since *all* the rabbits from month n will have produced a pair of offspring, in addition to the b pairs of rabbits alive in month $n + 1$. The new rabbits in month $n + 1$ will not have produced any offspring in month $n + 2$. The Fibonacci sequence is thus specified by:

$$F(1) = 1, \quad F(2) = 1 \quad \text{and} \quad F(n + 2) = F(n + 1) + F(n)$$

Alternatively, the Fibonacci sequence can also be specified by:

$$F(1) = 0, \quad F(2) = 1 \quad \text{and} \quad F(n + 2) = F(n + 1) + F(n)$$

which describes the sequence of new-born pairs of rabbits, and is the first sequence with each term shifted back one, given there are no new-born pairs at the start of the first month. A very accessible introduction to Fibonacci sequences and their applications can be found in Hammel-Garland *Fascinating Fibonaccis: Mystery and Magic in Numbers* (1987). Technology such as the CAS *Mathematica* can be readily used to generate terms of Fibonacci sequences and draw their graphs. The following recursion is different for the second order difference equation, since it is not based on iteration of a fixed point.

```
f[n_] := f[n - 1] + f[n - 2];f[0] = f[1] = 1
Table[f[n],{n,0,10}]
{1,1,2,3,5,8,13,21,34,55,89}
```

The corresponding graph is shown in Figure 5.5.

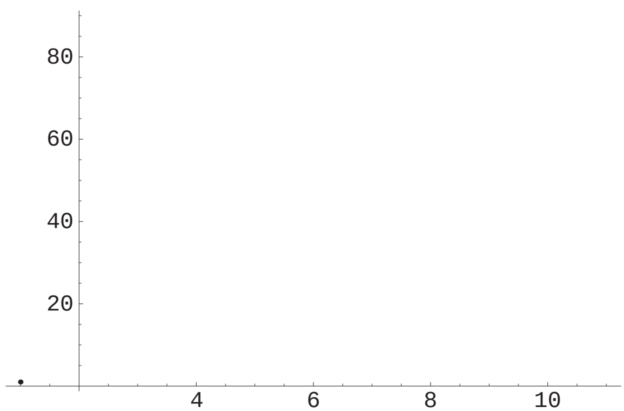


Figure 5.5: Graph of the first few terms of the Fibonacci sequence

The terms of the Fibonacci sequence are called the *Fibonacci numbers*. Just as CAS have built in applications for solving differential equations, they often also have similar applications for solving difference equations (recurrence relations). *Mathematica* has the discrete mathematics application **RSolve** for this purpose.

This application can be used to obtain the explicit formula for the Fibonacci sequence:

```
RSolve[{f[n] == f[n - 1] + f[n - 2], f[0] == f[1] == 1}, f[n], n]
```

$$\left\{\left\{f[n] \to \frac{\left(\frac{1}{2}(1-\sqrt{5})\right)^{n}\left(-1+\sqrt{5}\right)+2^{-n}\left(1+\sqrt{5}\right)^{1+n}}{2\sqrt{5}}\right\}\right\}$$

More generally, if the initial terms of *a* Fibonacci-type sequence are arbitrary integers *a* and *b* then this new sequence is called a Lucas sequence and is specified by:

$$f(1) = a \quad f(2) = b \quad \text{and} \quad f(n+2) = f(n+1) + f(n)$$

which can also be expressed in terms of *the* Fibonacci sequence by:

$$f(n+2) = bF(n+1) + aF(n)$$

A sequence of particular interest is the *Lucas* sequence {1, 3, 4, 7, 11 …} whose terms are called Lucas numbers. Edouard Lucas was a 19th-century French mathematician who studied recursion and sequences, and gave the Fibonacci sequence its name. The Lucas and Fibonacci numbers are linked by the relation $L(n) = F(n+1) + F(n-1)$. They are both connected to the golden ratio $\varphi = \frac{1+\sqrt{5}}{2}$ by the relationship:

$$\varphi^{n} = \frac{L(n) + \sqrt{5}F(n)}{2}$$

Indeed, these connections lead to an explicit (closed form) formula for the Fibonacci sequence.

Consider a line segment AC, with point B on the line segment between points A and C such that $AB > BC$, and $AB = 1$ as shown in Figure 5.6.

Figure 5.6: Geometric definition of the golden ratio

The golden ratio (section or mean), φ, is determined by the relation that $AC : AB = AB : BC$.

It occurs in several contexts in plant biology (usually through spiral arrangements of plant segment which are growing or packed together, such as in seed-heads) and is considered to correspond to a proportion that human beings find aesthetically pleasing in design contexts. Some authors also argue that the golden ratio can be found in architecture across cultures throughout history; however, like many hypotheses and conjectures, the validity of these claims this has been contested in recent analysis and commentary.

Let $AC = x$, then the requirement that $AC : AB = AB : BC$ is equivalent to:

$$\frac{x}{1} = \frac{1}{x-1}$$

which leads to the quadratic equation:

$$x^2 - x - 1 = 0$$

with roots $\frac{1-\sqrt{5}}{2}$ and $\frac{1+\sqrt{5}}{2}$. It is interesting to note that the *sum* of these two roots is 1, their *difference* is $\sqrt{5}$ and their *product* is –1. The positive root value is assigned to φ, consistent with its geometric interpretation of lengths as positive quantities. The other (negative) root has the exact value 1 – φ. Their numerical approximations correct to 3 decimal places are respectively 1.618 and –0.618.

It may be somewhat surprising to observe that the ratio of consecutive Fibonacci numbers $F(n + 1)$: $F(n)$ converges to φ as $n \to \infty$. The following list shows this ratio for the first 20 values of n, correct to 10 decimal places:

{1.0000000000, 2.0000000000, 1.5000000000, 1.6666666667, 1.6000000000, 1.6250000000, 1.6153846154, 1.6190476190, 1.6176470588, 1.6181818182, 1.6179775281, 1.6180555556, 1.6180257511, 1.6180371353, 1.6180327869, 1.6180344478, 1.6180338134, 1.6180340557, 1.6180339632}

To explain why this is the case, assume the limiting value of this ratio exists and is x, then taking the limit as $n \to \infty$ gives :

$$\begin{aligned} x &= \lim_{n\to\infty} \frac{F(n+1)}{F(n)} \\ &= \lim_{n\to\infty} \frac{F(n)+F(n-1)}{F(n)} \\ &= 1 + \lim_{n\to\infty} \frac{F(n-1)}{F(n)} \\ &= 1 + \lim_{n\to\infty} \frac{1}{\frac{F(n)}{F(n-1)}} \\ &= 1 + \frac{1}{x} \end{aligned}$$

which yields the same quadratic equation and roots as previously, for the limiting value. If this equation is written in the equivalent forms $x^2 = x + 1$ and

$x^{n+2} = x^{n+1} + x^n$ then the functions defined by $p(n) = \varphi^n$ and $q(n) = (1 - \varphi)^n$ as well as linear combinations $ap(n) + bq(n)$ of these satisfy the functional equation $f(n + 2) = f(n + 1) + f(n)$.

Determining the values of a and b so that the initial conditions for the Fibonacci sequence are satisfied gives:

$$F(n) = \frac{\varphi^n}{\sqrt{5}} - \frac{(1-\varphi)^n}{\sqrt{5}} = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{(\sqrt{5}) \times 2^n}$$

As n becomes increasingly large, the second term tends to zero, so the sequence of Fibonacci numbers converges on the nearest integer values of $\frac{\varphi^n}{\sqrt{5}}$.

The explicit rule for the Fibonacci sequence is a surprise: a formula involving powers of surd expressions that yields a positive integer when evaluated.

It is also possible to use matrices to compute arbitrary sub-sequences of length three of the Fibonacci sequence, that is, for a given n, compute the sub-sequence $\{F(n - 1), F(n), F(n + 1)\}$.

Consider the 2×2 matrix:

$$F = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Then the nth power of F is the matrix:

$$F^n = \begin{bmatrix} F(n+1) & F(n) \\ F(n) & F(n-1) \end{bmatrix}$$

from which the required sub-sequence can be listed. There is a nice straightforward proof of this result using *mathematical induction*. That is:

- show the result is true for the initial value (by inspection or simple calculation)
- show that if it is true for n then it is true for $n + 1$ (assume it is true for n and deduce from this that it is true for $n + 1$, by using established properties of matrices and the Fibonacci sequence)

When $n = 1$, the sub-sequence generated is $\{0, 1, 1\}$, the first three terms of the Fibonacci sequence starting from 0.

Assume $F^n = \begin{bmatrix} F(n+1) & F(n) \\ F(n) & F(n-1) \end{bmatrix}$ then:

$$F^{n+1} = F^n \times F = \begin{bmatrix} F(n+1) & F(n) \\ F(n) & F(n-1) \end{bmatrix} \times \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Hence $F^{n+1} = \begin{bmatrix} F(n+1)+F(n) & F(n+1) \\ F(n)+F(n-1) & F(n) \end{bmatrix}$

From the recursive definition of the Fibonacci sequence:
$F(n + 2) = F(n + 1) + F(n)$ and $F(n + 1) = F(n) + F(n - 1)$, so this is the same as:

$$F^{n+1} = \begin{bmatrix} F(n+2) & F(n+1) \\ F(n+1) & F(n) \end{bmatrix}$$

as required. The benefit of this approach is that it provides an explicit formula that computes the Fibonacci sub-sequences using integers. The following shows an implementation of this computation using the CAS *Mathematica*:

$$\mathbf{F} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

```
; f[n _] := MatrixPower[F, n] / /MatrixForm
Table [f[n], {n, 1, 10}] / /TableForm
```

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$$
$$\begin{pmatrix} 8 & 5 \\ 5 & 3 \end{pmatrix}$$
$$\begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix}$$
$$\begin{pmatrix} 21 & 13 \\ 13 & 8 \end{pmatrix}$$
$$\begin{pmatrix} 34 & 21 \\ 21 & 13 \end{pmatrix}$$
$$\begin{pmatrix} 55 & 34 \\ 34 & 21 \end{pmatrix}$$
$$\begin{pmatrix} 89 & 55 \\ 55 & 34 \end{pmatrix}$$

Various CAS list and data structure manipulation functions could then be used to extract the values of the Fibonacci sequence:

$$F = \{1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89 \ldots\}$$

Hoggatt (1969) gives an excellent introduction to Fibonacci sequences and Lucas numbers, and their relationship to the golden ratio.

THE LOGISTIC FUNCTION

The analysis of population models provides an ideal opportunity for students to engage in mathematical analysis, supported by graphical and numerical computation, where explicit solutions to the corresponding functional equations are not the most suitable form of representation to support such analysis. A simple model for population growth is the percentage increase, or *exponential* model. This model applies where there is effectively good access to resources, and a population increases over a fixed time interval (such as an hour, a year or a generation, depending on the context) by a fixed percentage of the previous population.

Thus, if the fixed percentage increase is 10%, then from the nth time interval to the $(n + 1)$th the population is 100% + 10% =110% of its previous value. Since 110% = 1.1 as a decimal, this can be expressed more concisely as a recurrence relation, where, if $p(n)$ represents the population at the (beginning of the) nth time interval:

$$p(n + 1) = 1.1 \times p(n) \quad \text{starting from some initial population value } p(0)$$

More generally, for a growth factor r, where $r \in R$ and $r > 1$ then $p(n + 1) = r \times p(n)$ for exponential growth.

While the exponential growth model can be used to describe the initial stages of growth of many populations, resources are usually not unlimited, and there are often other factors which affect the growth of a population.

A simple modification to the exponential growth model is to make the growth factor r a linear function of $p(n)$ such that $r = a(b - p(n))$. The parameter b can be thought of as representing some kind of ideal upper limit for a population in the given context, thus $p(n) < b$, while the parameter $a \in R^+$ plays a role somewhat like the original constant r in the exponential model. In practice, both a and b would be determined using experimental or historical data from the context under consideration.

This is called the *logistic* population model, and is a non-linear first order difference equation. Clearly, when $p(n)$ is small, $r \approx ab$, and growth is almost exponential. Conversely, as $p(n) \rightarrow b$, then $r \rightarrow 0$ and growth effectively ceases.

A population is defined be *stable* if the population does not change from one time interval to the next time interval. That is, an *equilibrium* value occurs when $p_{eq}(n+1) = p_{eq}(n)$, and hence $r = 1$.

Now, $r = 1 \quad \Rightarrow \quad a(b - p_{eq}(n)) = 1$

$\qquad\qquad\quad \Rightarrow \quad p_{eq}(n) = b - \frac{1}{a}$

For example, assume that two breeding pairs of a certain animal are released into a new and isolated environment with initially plentiful but ultimately limited resources. Initially the population might double each time interval, that is $r = 2$. Suppose that previous data suggests that the environment the animals have been released into can support an idealised maximum population of around $b = 500$ animals.

Then the value of a can be determined, since $r = a(b - p(n))$, the data gives $2 = a(500 - 4)$, that is $a = \frac{2}{496} \approx 0.004$. The value for the stable population can now be calculated as $p_{eq}(n) = b - \frac{1}{a} \approx 250$ animals. This can be checked by explicitly generating a sequence of values for this population:

```
f[n_] := 0.004n(500 - n)
NetList[f,4,10]
{4,7.936,15.6201,30.2642,56.8647,100.795,160.952,
218.282,245.976,249.935,250.}
```

and illustrated by drawing the corresponding population graph as shown in Figure 5.7.

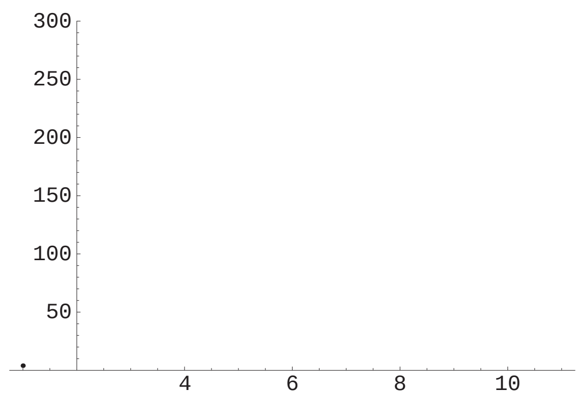


Figure 5.7: Logistic population model graph

Analysis of this type of situation can be carried out more generally by taking b to be 1 and restricting $p(n)$ between 0 and 1. Thus $p(n)$ could be interpreted as a kind of *proportion* of an ideal population, with minimum value 0 and maximum value 1. In this formulation the logistic population difference equation model is now of the form $p(n+1) = a(1-p(n))\,p(n)$ with $a > 0$ as the only parameter, and the behaviour of this system can be explored in terms of this parameter. As before, there will be an equilibrium value at $p_{eq}(n) = 1 - \frac{1}{a}$ and $a > 1$ (to ensure $p(n) > 0$).

For example, suppose $p(0) = 0.3$ and a = 2.5, then $p_{eq}(n) = 0.6$, as can be seen from Figure 5.8.

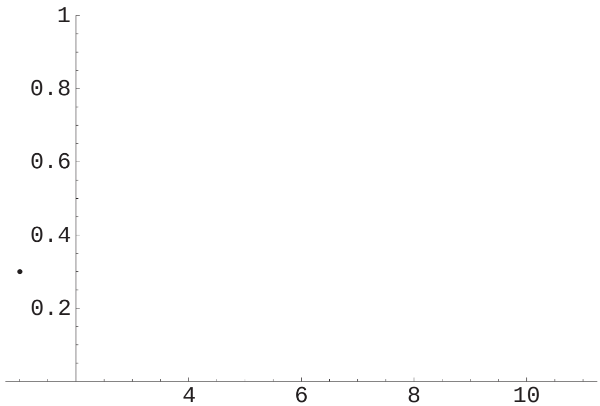


Figure 5.8: Logistic population model graph $a = 2.5$

For a just a little larger than 3, for example $a = 3.17$, *stable oscillations* occur, as illustrated in Figure 5.9.

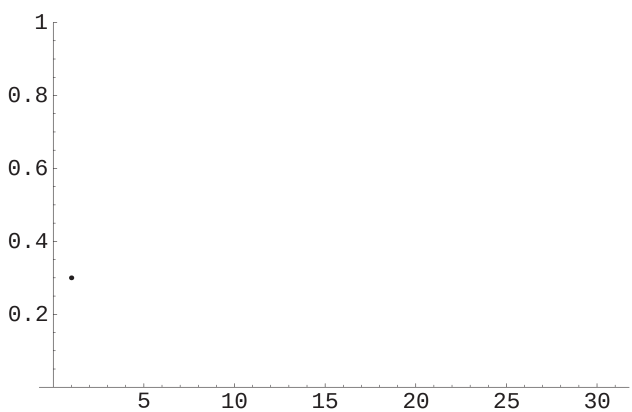


Figure 5.9: Logistic population model graph $a = 3.17$

And for a about 3.57, *chaotic* behaviour occurs, as illustrated in Figure 5.10.

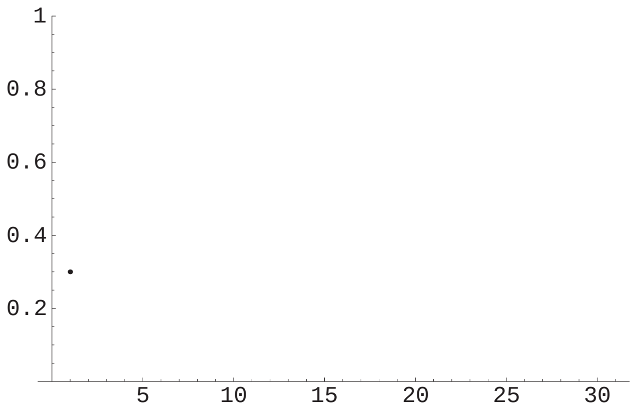


Figure 5.10: Logistic population model graph a = 3.57

This behaviour is fascinating for mathematicians since it comes from a *deterministic* model (a non-linear difference equation), but produces random-like data (see Potts 1985; Gleick 1987).

SUMMARY

- A sequence is a function f with domain N^+, and is usually described by listing its range in natural order $\{f(1), f(2), f(3), f(4) \ldots\}$.
- Sequences can be generated either by explicit rule '$f(n) =$' in terms of n, or recursively where $f(n + 1)$ is described in terms of $f(n)$ and possibly other preceding terms.
- A difference equation or recurrence relation, is a solution to a functional equation over domain N^+ that is recursively defined.
- An arithmetic sequence is defined recursively by $f(1) = a$ and $f(n + 1) = f(n) + d$. Its explicit rule form is $f(n) = a + (n - 1)d = dn + (a - d)$.
- A geometric sequence is defined recursively by $f(1) = a$ and $f(n + 1) = r \times f(n)$. Its explicit rule form is $f(n) = a \times r^{n-1} = \frac{a}{r} \times r^n$.
- A first order linear difference equation is defined recursively by $f(n + 1) = af(n) + b$ where $n \in N^+$ and $a, b \in R$. Its explicit rule form is $f(n) = an^{-1} f(1) + \dfrac{b(a^{n-1} - 1)}{a - 1}$, where $a \neq 1$.
- The *constant* sequence is a special case of a first order linear difference equation where $a = 0$; the *arithmetic* sequence is a special case of a first order linear difference equation where $a = 1$; and the *geometric* sequence is a special case of a first order linear difference equation where $b = 0$.

- A quadratic function $f: N \to R$, $f(x) = ax^2 + bx + c$ is a solution to the second order linear difference equation: $f(n+2) = 2f(n+1) - f(n) + k$ where k is the constant second difference from the corresponding finite difference table and $k = 2a$.
- The Fibonacci function $f: N \to R$,

  $$F(n) = \frac{\varphi^n}{\sqrt{5}} - \frac{(1-\varphi)^n}{\sqrt{5}} = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{(\sqrt{5}) \times 2^n}$$ where $F(0) = 1$ and $F(1) = 1$ is a solution to the difference equation $f(n+2) = f(n+1) + f(n)$. More general solutions with other initial conditions are called Lucas sequences.
- The logistic function $f(n+1) = a(1 - f(n))\, f(n)$ where $0 < f(n) < 1$ is an example of a first order non-linear difference equation, and is used to model various populations. For different values of the parameter $a > 0$ it exhibits convergent, oscillating and chaotic behaviour.

STUDENT ACTIVITY 5.1

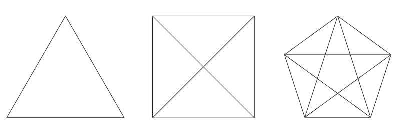
For the purposes of determining its length, a spiral wound tape can be thought of as a series of concentric circles, with the radii of consecutive circles differing by a fixed amount equal to the thickness of the tape. Describe how this can be used to estimate the length of the tape, and apply this method to a real roll of tape, such as plastic tape.

STUDENT ACTIVITY 5.2

A table-tennis ball is dropped from a height of one metre and allowed to bounce repeatedly until it stops. If it is assumed that the ball bounces to the same proportion of its previous maximum height on each bounce, find a good estimate for the total distance it travels before it comes to rest.

STUDENT ACTIVITY 5.3

As shown in the diagrams below, a triangle has no diagonals, a square has two diagonals and a pentagon has five diagonals.



Find a formula for the number of diagonal of an *n*-sided polygon.

STUDENT ACTIVITY 5.4

A $50 000 loan is taken out to buy a new family car. If the interest is 16% annually and calculated quarterly, what is the fixed amount that needs to be paid each quarter (that is, the identical payment made each quarter over the life of the loan) if the loan is to be paid off over 10 years?

STUDENT ACTIVITY 5.5

Investigate how the terms of the Fibonacci sequence are related to the coefficients of binomial expansion in Pascal's triangle (Hint: write out Pascal's triangle as left-justified text, and then draw in diagonal lines from the top right down to the bottom left).

Using 1 cm grid graph paper, for each of the first few terms of the Fibonacci sequence cut out a square with side length corresponding to that term. Use these to construct a nested series of rectangles with side lengths corresponding to the values of the first n terms of the Fibonacci sequence.

STUDENT ACTIVITY 5.6

Investigate the graphical behaviour of the logistic model difference equation for different values of the parameter a, and different initial values of $p(0)$.

References

Cohen, SJ 2002, *Computer algebra and symbolic computation*, Peters, Natick MA.

Hammel-Garland, 1987, *Fascinating Fibonaccis: Mystery and magic in numbers*, Dale Seymour, Palo Alto CA.

Hodgson, B & Leigh-Lancaster, D 1990, *Space and number*, Jacaranda Press, Brisbane.

Hoggatt, VE 1969, *Fibonacci and Lucas numbers*, Houghton Mifflin, New York.

Potts, RB 1985, Discrete Mathematics, Plenary presentation, *Proceedings of the Fifth International Congress on Mathematics Education*. M Carss (ed.), Birkhäuser, Boston.

Websites

http://www.mcs.surrey.ac.uk/Personal/R.Knott/Fibonacci/fib.html – Ron Knott, University of Surrey

This website contains and extensive collection Fibonacci sequence related material, including many diagrams and pictures. See also the related website by the same author: http://www.mcs.surrey.ac.uk/Personal/R.Knott/Fibonacci/fibnat.html

http://mathworld.wolfram.com/RecurrenceEquation.html – Wolfram Research

Online encyclopaedia by the developers of *Mathematica*. An extensive and rigorous mathematical encyclopaedia, with extensive cross-references.

http://chaos.phy.ohiou.edu/~thomas/chaos/logistic.html - Ohio State University

Provides an overview of the logistic model and its behaviour.

CHAPTER 6

CURRICULUM CONNECTIONS

Different school systems and educational jurisdictions have particular features in their senior secondary mathematics curricula that have been developed over decades, and even centuries in some cases, to meet the historical and contemporary educational needs of their cultures and societies. When these curricula are reviewed, it often the case that this includes a process of benchmarking with respect to corresponding curricula in other systems and jurisdictions. This may be in a local, county, state, national or international context.

Over the past few decades, particularly in conjunction with renewed interest in comparative international assessments (such as TIMSS and PISA OECD), curriculum benchmarking has been employed extensively by educational authorities and ministries. Such benchmarking reveals much that is common in curriculum design and purpose in senior secondary mathematics courses around the world. Some key design constructs that can are used to characterise the nature of senior secondary mathematics courses are:

- content (areas of study, topics, strands)
- aspects of working mathematically (concepts, skills and processes, numerical, graphical, analytical, problem-solving, modelling, investigative, computation and proof)
- the use of technology, and when it is permitted, required or restricted (calculators, spreadsheets, statistical software, dynamic geometry software, computer algebra systems)
- the nature of related assessments (examinations, school based and the relationship between these)
- the relationship between the final year subjects and previous years in terms of the acquisition of important mathematical background (assumed knowledge and skills, competencies, prerequisites and the like)
- the amount and nature of prescribed material within the course (completely prescribed, unitised, modularised, core plus options)

- the amount of in-class (prescribed) and out-of-class (advised) time that students are expected to spend on completion of the course

In broad terms, it is possible to characterise four main sorts of senior secondary mathematics courses.

Type 1

Courses designed to consolidate and develop the foundation and *numeracy* skills of students with respect to the practical application of mathematics in other areas of study. These often have a *thematic* basis for course implementation.

Type 2

Courses designed to provide a *general* mathematical background for students proceeding to employment or further study with a *numerical* emphasis, and likely to draw strongly on *data analysis* and *discrete* mathematics. Such courses typically do not contain any calculus material, or only basic calculus material, related to the application of average and instantaneous rates of change. They may include, for example, business-related mathematics, linear programming, network theory, sequences, series and difference equations, practical applications of matrices and the like.

Type 3

Courses designed to provide a sound foundation in function, coordinate geometry, algebra, *calculus* and possibly probability with an *analytical* emphasis. These courses develop mathematical content to support further studies in mathematics, the sciences and sometimes economics.

Type 4

Courses designed to provide an *advanced* or *specialist* background in mathematics. These courses have a *strong analytical* emphasis and often incorporate a focus on mathematical *proof*. They typically include complex numbers, vectors, theoretical applications of matrices (for example transformations of the plane), higher level calculus (integration techniques, differential equations), kinematics and dynamics. In many cases Type 4 courses assume that students have previous or concurrent enrolment in a Type 3 course, or subsume them.

Table 6.1 provides a mapping in terms of curriculum connections between the chapters of this book, the four types of course identified above, and the courses currently (2005) offered in various Australian states and territories. As this book is a *teacher* resource, these connections are with respect to the usefulness of material from the chapters in terms of mathematical background of relevance, rather than direct mapping to curriculum content, or syllabi, in a particular state or territory.

Table 6.1: Curriculum connections for senior secondary final year mathematics courses in Australia

State or territory	*Type of course*	*Relevant chapters*
Victoria	2: Further Mathematics	5
	3: Mathematical Methods/ Mathematical Methods(CAS)	1, 2, 3 and 4
	4: Specialist Mathematics	1, 2, 3 and 4
New South Wales	2: General Mathematics	5
	3: Mathematics and Mathematics Extension 1	1, 2, 3 and 4
	4: Mathematics Extension 2	1, 2, 3 and 4
Queensland	2: Mathematics A	5
	3: Mathematics B	1, 2, 3 and 4
	4: Mathematics C	1, 2, 3 and 4
South Australia/ Northern Territory	2: Mathematical Applications	5
	3: Mathematical Methods/Mathematical Studies	1, 2, 3 and 4
	4: Specialist Mathematics	1, 2, 3 and 4
Western Australia	2: Discrete Mathematics	5
	3: Applicable Mathematics	1, 2, 3 and 4
	4: Calculus	1, 2, 3 and 4
Tasmania	2: Mathematics Applied	5
	3: Mathematics Methods	1, 2, 3 and 4
	4: Mathematics Specialised	1, 2, 3 and 4

Table 6.2 provides a mapping in terms of curriculum connections between the chapters of this book, the four types of course identified above, and some of the courses currently (2005) offered in various English-speaking jurisdictions from around the world. Again, as this book is a *teacher* resource, these connections from the usefulness of material from the chapters in terms of mathematical background of relevance, rather than direct mapping to curriculum content, or syllabuses, in a particular jurisdiction.

Table 6.2: Curriculum connections for senior secondary final year mathematics courses in some jurisdictions around the world

State or territory	*Type of course*	*Relevant chapters*
College Board US	3: Advanced Placement Calculus AB	1, 2, 3 and 4
	4: Advanced Placement Calculus BC	1, 2, 3 and 4
International Baccalaureate Organisation (IBO)	3: Mathematics SL	1, 2, 3 and 4
	4: Mathematics HL	1, 2, 3 and 4
UK	3: AS Mathematics	1, 2, 3 and 4
	4: Advanced level	1, 2, 3 and 4

Content from the chapters of the book may be mapped explicitly to topics within particular courses, and teachers will perhaps find it useful to informally make these more specific connections in terms of their intended implementation of a given course of interest to them.

References

The following are the 2005 website addresses of Australian state and territory curriculum and assessment authorities, boards and councils. These include various teacher reference and support materials for curriculum and assessment.

The Senior Secondary Assessment Board of South Australia (SSABSA)
http://www.ssabsa.sa.edu.au/

The Victorian Curriculum and Assessment Authority (VCAA)
http://www.vcaa.vic.edu.au/

The Tasmanian Qualifications Authority (TQA)
http://www.tqa.tas.gov.au/

The Queensland Studies Authority (QSA)
http://www.qsa.qld.edu.au/

The Board of Studies New South Wales (BOS)
http://www.boardofstudies.nsw.edu.au/

The Australian Capital Territory Board of Senior Secondary Studies (ACTBSSS)
http://www.decs.act.gov.au/bsss/welcome.htm

The Curriculum Council Western Australia
http://www.curriculum.wa.edu.au/

The following are the 2005 website addresses of various international and overseas curriculum and assessment authorities, boards, councils and organisations:

College Board US Advanced Placement (AP) Calculus
http://www.collegeboard.com/student/testing/ap/sub_calab.html?calcab

International Baccalaureate Organisation (IBO)
http://www.ibo.org/ibo/index.cfm

Qualifications and Curriculum Authority (QCA) UK
http://www.qca.org.uk/

OECD Program for International Student Assessment (PISA)
http://www.pisa.oecd.org

Trends in International Mathematics and Science Study (TIMSS)
http://nces.ed.gov/timss/

CHAPTER 7

SOLUTION NOTES TO STUDENT ACTIVITIES

Student activity 1.1

a ‘is older than’ is *not* an equivalence relation as it is not *reflexive*, that is, a person is not older than themselves (it is not symmetric either, but it is transitive)

b ‘$\leq$’ is *not* an equivalence relation on R as it is not *symmetric*, that is $x \leq y$ does not imply $y \leq x$, for example $3 \leq 10$ does not imply $10 \leq 3$ (it is, however, reflexive and transitive)

c ‘$\subseteq$’ is not an equivalence relation on subsets of the roman alphabet as it is not *symmetric*, that is $x \subseteq y$ does not imply $y \subseteq x$, for example, {vowels} $\subseteq$ the roman alphabet but the roman alphabet $\subseteq$ {vowels}. However, it is reflexive and transitive.

d ‘divides’ on N is *not* an equivalent relation, as it is not *symmetric*, that is x divides y does not imply y divides x, for example 3 divides 12 does not imply 12 divides 3 (however, 3 divides itself, and also divides any number that 12 divides, since it is a factor of 12).

Examples of other relations that are equivalence relations are ‘parallel’ and ‘congruent’ in geometry and ‘equivalent-fraction’ in number.

Student activity 1.2

Students should identify this for the technology they are using. For the CAS *Mathematica* these are summarised in the following table:

Operation	*Example*
Assign a constant	`a = 3`
Denote a variable	Any symbol or expression such as `m`, `x1`, `yyy` or `area` can be used as a variable. The expression `x_` denotes a free variable in the left-hand side of the definition of a function. This means that any numerical value, symbolic expression or combination of these can be substituted for an occurrence of `x` in the rule of the function.
Define the rule of a function	`f[x_]:= x^2 - x` defines `f` as a function of the free variable `x`.
Solve an equation	`Solve[f[x]== 0, x]` solves the given equation for x. The use of `==` does not assign the value `0` to `f[x]` as the expression `f[x]= 0` would, but finds the value(s) of `x` which makes the given statement true.
Clearing definitions or assignments	`Clear[f]` and `Clear[a]` clears the definition of the function `f` and the constant `a` respectively.
Specifying conditions or constraints	`f[n_]:=n!/;n >0` defines `f` as the factorial function, given that '`/;`' `n` is positive.

Depending on the context, there are often several ways to implement a process using a given technology, and the benefits and limitation of these are typically learnt over an extended period use.

Students can compare the relevant approaches for their CAS with those listed for *Mathematica* above. This activity can be used to alert student to generic and idiosyncratic aspects of different technologies as they are constructed to model and implement various aspects of mathematics.

Student activity 1.3

a ☐ = 40. Some students will obtain this value by computation; others will note that since 24 is 4 less than 28, then ☐ must be 4 more than 36, that is ☐ = 40.

b ☐ = 106. Again either computation or balancing can be used.

c $n - 1 + 7 = n + 5 + 1$. Students who use the balancing approach to parts a and b will often be able to generalise to this result. Some students will identify a value of n or some values of n.

d Some students will attempt to solve as an equation for n, and reduce to an inconsistent form, others will see the equivalent form with the left hand side as $n - 3 + n - 3$ and the right hand side as $n - 4 + n - 3$, and hence detect the inconsistency.

Student activity 1.4

a $x = \frac{7}{2}$

b *all* real values of x, this is an algebraic identity, and reduces to $0 = 0$, a statement which is always *true* independent of the value of x.

c *no* real values of x, this is a mathematical contradiction, and reduces to $0 = 1$, a statement which is always *false* independent of the value of x.

Student activity 1.5

a If $f = g$ then $f(x) = g(x) + c$, where c is a real constant implies $f'(x) = g'(x)$. That is, if two functions differ by a constant, then their derivatives are equal.

b If $f = g$ then $\int f(x)\,dx = \int g(x)\,dx + c$, where c is a real constant. That is, if two functions are equal then their antiderivatives will differ by a constant (which may be zero).

c If $f = g$, $\int f(x)\,dx = F(x)$ and $\int g(x)\,dx = G(x)$ then $F(x) = G(x) + c$ where c is an arbitrary real constant. Thus:

$$\int_a^b f(x)dx = F(b) - F(a) = (G(b) + c) - (G(a) + c) = G(b) - G(a)$$
$$= \int_a^b g(x)dx$$

Student activity 1.6

a If the quadratic function is defined by $f: R \to R$, $f(x) = ax^2 + bx + c$, then the vertex of its graph occurs when $f'(x) = 0 \Rightarrow 2ax + b = 0 \Rightarrow x = -\frac{b}{2a}$. To establish that this is a line for reflection symmetry, it needs to be shown that $f(\frac{-b}{2a} - h) = f(\frac{-b}{2a} + h)$ for any positive real valued h. This follows readily from recognition that the completed square form for f will be $a(x + (\frac{b}{2a}))^2 + k$ for some real constant k.

b If the cubic function is defined by $f: R \to R$, $f(x) = ax^3 + bx^2 + cx + d$, then the point of inflection of its graph occurs when $f''(x) = 0 \Rightarrow 6ax + 2b = 0 \Rightarrow x = \frac{-b}{3a}$, and lies on the horizontal line $y = f(\frac{-b}{3a}) = k$. The point $(\frac{-b}{3a}, k)$ is a point of half-turn rotational symmetry for the graph of f if, for any positive real valued h, then $f(\frac{-b}{3a} + h) - k = -(f(\frac{-b}{3a} - h) - k)$. This can be shown algebraically by extended manipulation; however, CAS will verify the result quickly. Unfortunately, there is no simple argument based on a 'completed cube' form as no such general form exists for cubic polynomial functions (the existence of such a form would imply that the graph of *any* cubic function is a simple transformation of the graph of $y = x^3$, which is clearly not the case for any cubic function whose graph has a *local maximum* and *local minimum*, or no stationary point of inflection).

Student activity 2.1

The purpose of this activity is to get students to think about appropriate graphing windows when using technology to illustrate the essential features of a function through its graph over a suitable subset of its natural domain. Key points to consider are:

- self-similarity
- asymptotic behaviour
- symmetry
- whether the function is increasing or decreasing over a given interval or has any stationary points

Student activity 2.2

b, *d*, *e*, *h*, *j* and *m*.

Student activity 2.3

a, *c*, *g* and *i*

Student activity 2.4

The constant function $f: R \to R, f(x) = 0$.

Student activity 2.5

a None of the examples used in this chapter are self-inverses.

b The real-valued functions with rules $f(x) = x$ and $f(x) = \frac{1}{x}$ are self-inverses.

Student activity 2.6

a When the product of two real numbers is equal to their sum, which gives pairs (x, y) of the form $(x, \frac{x}{x-1})$ where x is a real numbers and $x \neq 1$. For integers: $x = y = 0$ or $x = y = 2$, or an integer and rational example is $(-1, \frac{1}{2})$.

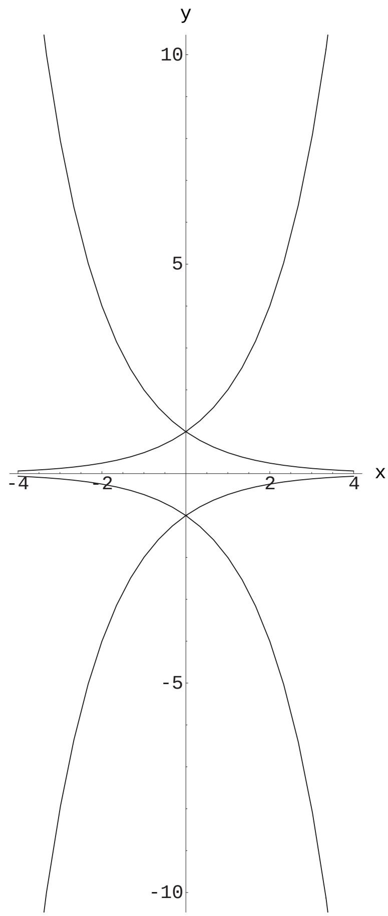
b When the square of the product of two real numbers is equal to the sum of their squares, which gives pairs (x, y) of the form $(x, \pm\frac{x}{\sqrt{x^2-1}})$ where $|x| > 1$.

Student activity 2.7

Interpret the function in terms of an area relation.

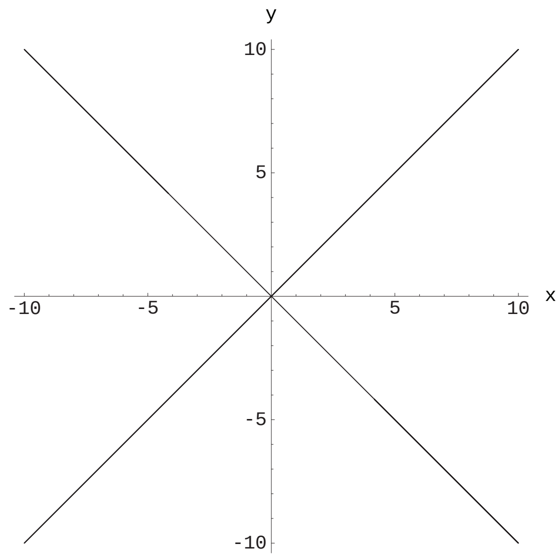
Student activity 3.1

a If $f(x) = 2^x$ then $f(-x) = 2^{-x} = \frac{1}{2^x} = (\frac{1}{2})^x$. The graphs of $f(x)$, $f(-x)$, $-f(x)$ and $-f(-x)$ are shown below:



Clearly $f(x) \neq f(-x)$, so f is not an even function (it is not symmetrical by reflection in the vertical axis); similarly $-f(x) \neq f(-x)$, so f is not an odd function (it is not symmetrical by half-turn rotation about the origin).

b The graph of the relation $y^2 = x^2$ is shown below, and is symmetrical both with respect to reflection in the vertical axis and half turn rotation about the origin, thus it is both an even and an odd relation.



Student activity 3.2

Only $f(x) = 4x$ is a solution.

Student activity 3.3

Only $f(x) = \sin(x)$ and $f(x) = \cos(x)$ are solutions.

Student activity 3.4

a $f(x) = \pm 1$
b $f(x) = x$, $f(x) = -x$, $f(x) = \frac{1}{x}$, $f(x) = -\frac{1}{x}$
c $f(x) = x$, $f(x) = -x$

Student activity 3.5

The functional equation $f(-x) = -f(x)$ can be seen as a special case of the functional equation $f(kx) = kf(x)$ where $k = -1$. For real-valued differentiable functions, $f(-x) = -f(x)$ has all odd functions as solutions, including power functions of the form $y = kx^n$ where n is an odd integer; while $f(kx) = kf(x)$ has functions of the form $y = ax$, where a is a real constant as solutions.

Student activity 3.6

a *Constant*: let $f(x) = k$, where k is a real constant. Then $f(x)\,f(-x) = f(x^2)$ implies $k^2 = k$, so $f(x) = 0$ or $f(x) = 1$.
b *Linear*: let $f(x) = ax + b$, where $a \neq 0$, then $f(x)\,f(-x) = f(x^2)$ implies $(ax + b)(\,-ax + b\,) = ax^2 + b$, so equating coefficents gives $-a^2 = a$ where $a \neq 0$ and $b^2 = b$. Hence $a = -1$ and $b = 0$ or 1.

c *Quadratic*: let $f(x) = ax^2 + bx + c$, where $a \neq 0$, then $f(x)\,f(-x) = f(x^2)$ implies
$(ax^2 + bx + c)(ax^2 - bx + c) = ax^4 + bx^2 + c$, so equating coefficients gives $a = 1, b = c = 0$; or $a = b = c = 1$; or $a = 1, b = -1, c = 0$; or $a = c = 1, b = -2$.

d Let $x^2 f(x) + f(1-x) = 2x - x^4$ then $f(1-x) = 2x - x^4 - x^2 f(x)$. Interchanging x with $1 - x$ in $x^2 f(x) + f(1 - x) = 2x - x^4$ gives
$(1-x)^2 f(1-x) + f(x) = 2(1-x) - (1-x)^4$. Substituting
$f(1-x) = 2x - x^4 - x^2 f(x)$ and simplifying gives $f(x) = 1 - x^2$.

Student activity 4.1

Only $f(x) = x$ is a solution.

Student activity 4.2

a $f(x - y) = a(x - y) = ax - ay = f(x) - f(y)$, graphically this corresponds to asserting that the vertical height from the horizontal axis to the graph of $f(x) = ax$ at $x - y$ is the same as the difference between the vertical height at x from the horizontal axis to the graph of $f(x)$ and the vertical height at y from the horizontal axis to the graph of $f(x)$

b If $f(x) = ax + b$ then $f(x_2 - x_1) = a(x_2 - x_1) + b$ while $f(x_2) - f(x_1) = a(x_2 - x_1)$. Thus in this case the functional value of the difference is the same as the difference of the functional values *plus* a vertical translation of b.

Student activity 4.3

a Let m and n be non-zero integers, then $g(x + y) = g(x) + g(y) \Rightarrow$
$\frac{1}{m+n} = \frac{1}{m} + \frac{1}{n} \Rightarrow m^2 + mn + n^2 = 0$. If m and n are the *same* sign this is not possible, since all terms will be non-zero and positive, hence
$m^2 + mn + n^2 > 0$. If m and n are of *opposite* sign then one of m^2 or n^2 is at least the same size as mn. As both m^2 and n^2 are non-zero this implies
$m^2 + n^2 - |mn| > 0$, so again this is not possible. Hence
$g(x + y) \neq g(x) + g(y)$ for integer x and y.

b $g(x + y) = (x + y)g(xy)$

Student activity 4.4

a $f(x) = ax + b$

b $f(x) = \cos(kx)$

c $f(x) = ax$, $f(x) = a\sin(kx)$

d $f(x + y) = \frac{f(x) + f(y)}{1 - f(x)f(y)}$

e $af(\sqrt{x^2 + y^2}) = f(x)\,f(y)$

Student activity 4.5

The first part requires the extension of the definition of an exponent to rational values, then the functional equations hold by virtue of the index laws. For the second part any *natural number* counter-example suffices, since $N \subset Q$.

Student activity 4.6

The natural logarithm is defined by the (integral) functional equation

$$\log_e(x) = \int_1^x \frac{1}{t}dt,\ t > 0.$$

If $f(x) = \int_1^x \frac{1}{t}dt$ this is clearly a function which has domain R^+ and is undefined at $x = 0$. By the properties of a definite integral, $f(1) = 0$.

By definition, $f(ab) = \int_1^{ab} \frac{1}{t}dt = \int_1^{a} \frac{1}{t}dt + \int_a^{ab} \frac{1}{t}dt$

Using the change of variable $t = au$, gives $\frac{dt}{du} = a$, $t = a$ gives $u = 1$ and $t = ab$ gives $u = b$, hence $f(ab) = \int_1^{a} \frac{1}{t}dt + \int_1^{b} \frac{1}{u}du = = f(a) + f(b)$.

From this it can be readily deduced by letting $b = \frac{1}{a}$ that $f(\frac{1}{a}) = -f(a)$ and $f(\frac{b}{a}) = f(b) - f(a)$. These functional equations characterise a logarithm function.

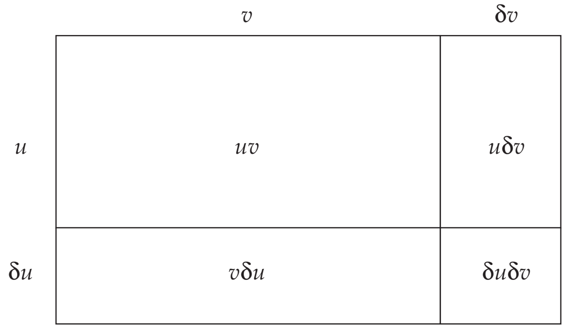
Student activity 4.7

Matrix multiplication is not, in general, commutative, that is, $XY \neq YX$. By definition $(XY)^2 = XYXY$. This will only be equal to X^2Y^2 if $XY = YX$ since the middle product could then be written as XY resulting in $XXYY = X^2Y^2$.

By definition of matrix inverse for multiplication, $(XY)^{-1}(XY) = I$. The product $(Y^{-1}X^{-1})(XY) = Y^{-1}(X^{-1}X)Y = Y^{-1}(I)Y = I$. As inverse are unique, this gives $(XY)^{-1} = Y^{-1}X^{-1}$.

Student activity 4.8

The product theorem for differentiation can be established from first principles using the limit definition of a derivative. It also has a simple geometric representation, since products can be represented geometrically as areas of rectangles. For example, let y, u and v all be functions of the real variable x, where $y = uv$. If a small change in the value of x is represented by δx, and consequent changes in y, u and v by δy, δu and δv respectively, then the change in the product function y can be identified with certain areas as shown below.



Using the diagram, it can be seen that:

$$\begin{aligned} y + \delta y &= (u + \delta u)(v + \delta v) \\ &= uv + u\delta v + v\delta u + \delta u\delta v \end{aligned}$$

Since $y = uv$, this is equivalent to:

$$\delta y = u\delta v + v\delta u + \delta u\delta v$$

Dividing through by δx gives:

$$\frac{\delta y}{\delta x} = u\frac{\delta v}{\delta x} + v\frac{\delta u}{\delta x} + \delta u\frac{\delta v}{\delta x}$$

taking the limit as $\delta x \to 0$ yields the result:

$$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$$

(since $\delta u \to 0$ as $\delta x \to 0$, the limiting value of the term $\delta u\frac{\delta v}{\delta x}$ is 0).

This is clearly *not* the same as the product of the derivatives of the original functions. A simple counter-example can be found by choosing $u = f(x) = x^2$ and $v = g(x) = x^3$.

Then $\frac{d(uv)}{dx} = 5x^4$ but $\frac{du}{dx}.\frac{dv}{dx} = 2x3x^2 = 6x^3$.

By writing $\frac{u}{v}$ as uv^{-1} and using the product and chain rules for differentiation the quotient rule can be obtained:

$$\frac{d(uv^{-1})}{dx} = u\frac{d(v^{-1})}{dx} + v^{-1}\frac{du}{dx} = u\frac{-1}{v^2}\frac{dv}{dx} + \frac{1}{v}\frac{du}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$$

and this is clearly *not* the same as the quotient of the derivatives of the original functions. Again, $u = f(x) = x^2$ and $v = g(x) = x^3$ provide a counter-example as:

$$\frac{d(uv^{-1})}{dx} = \frac{-1}{x^2} \text{ and } \frac{\frac{du}{dx}}{\frac{dv}{dx}} = \frac{2x}{3x^2} = \frac{2}{x}.$$

Student activity 4.9

Use $\tan(\theta + \varphi) = \frac{\sin(\theta + \varphi)}{\cos(\theta + \varphi)}$, expand the right hand side using the identities for $\sin(\theta + \varphi)$ and $\cos(\theta + \varphi)$, then divide each term by $\cos(\theta)\cos(\varphi)$ to obtain:

$$\tan(\theta + \varphi) = \frac{\tan(\theta) + \tan(\varphi)}{1 - \tan(\theta)\tan(\varphi)}$$

Let $\theta = \varphi$ in each of the original identities to obtain:

$\sin(2\theta) = 2\sin(\theta)\cos(\theta)$ $\quad\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$ $\frac{2\tan(\theta)}{1 - \tan^2(\theta)}$

Let $\varphi = -\theta$ in each of the original identities, and use the fact that sine is an odd function, cosine is an even function, and hence tangent is an odd function to obtain:

$$\sin(\theta - \varphi) = \sin(\theta)\cos(\varphi) - \cos(\theta)\sin(\varphi)$$
$$\cos(\theta - \varphi) = \cos(\theta)\cos(\varphi) + \sin(\theta)\sin(\varphi)$$
$$\tan(\theta - \varphi) = \frac{\tan(\theta) - \tan(\varphi)}{1 + \tan(\theta)\tan(\varphi)}$$

Student activity 5.1

Assume the roll of tape is in the form of an annulus (a circular ring) with inside and outside radii of r and R respectively. If the wound tape has k layers (each treated as a circle), then these layers have a common thickness of $d = \frac{R-r}{k}$. The length of each layer is then the circumference of the corresponding circle, which is 2π times its radius. The radii of the circles form an *arithmetic* sequence with first term r and constant difference d. The total length of the tape is then $2\pi \times$ sum of the radii (a similar problem is to estimate the length of the track on an old-fashioned vinyl record).

Student activity 5.2

If the ball bounces to, say 80% of its previous height on a given bounce, then this forms a *geometric* sequence of values: 1, $1 \times 0.8 = 0.8$, $1 \times 0.8 \times 0.8 = 0.64$ … and so on. As the common ratio $r = 0.8$ is such that $|r| < 1$, then the corresponding infinite geometric series converges with first term $a = 1$ and

$r = 0.8$. The value if the limiting sum is $\frac{a}{1-r}$, which in this case is $\frac{1}{0.2} = 5$ metres (in practice a 'standard' table-tennis ball often has a bounce factor of just less than 0.8 and bounces around 23 times before it stops).

Student activity 5.3

Continuing the geometric pattern develops a table of values:

Sides *s*	3	4	5	6	7	8
Diagonals *d*	0	2	5	9	14	20

Use of difference table gives a constant second difference of 1 so a quadratic polynomial function $d(s) = as^2 + bs + c$ will provide a suitable model, with $2a = 1$, so $a = \frac{1}{2}$. Solving $d(6) - d(4) = 7$ for b gives $b = -\frac{3}{2}$. Then solving $d(3) = 0$ for c gives $c = 0$, so $d(s) = \frac{1}{2}s^2 - \frac{3}{2}s$.

Alternatively, from the geometry of the situation, a polygon of s sides also has s vertices, and each vertex will be joined to $s - 3$ other vertices by a diagonal (not itself or either of its neighbours). Since the product $s \times (s - 3)$ counts the diagonal from a vertex to any other vertex *twice*, the number of diagonals is half this, that is $d(s) = \frac{1}{2} s \times (s - 3)$, which when expanded is algebraically equivalent to the earlier result (the geometric argument can also be adapted to deduce the form $d(s) = \frac{1}{2}s^2 - \frac{3}{2}s$ directly).

Student activity 5.4

Let $t(n)$ be the amount left to repay at the beginning of the nth quarter, and let p be the quarterly repayment required to pay off the loan over 10 years, with a rate of interest of 16% per annum. Then $t(1) = 50\,000$ and $t(n + 1) = 1.04t(n) - p$. After 10 years, or 40 quarters, the amount left to pay should be zero, so $t(41)$ = amount to pay at the *beginning* of the 41st quarter = 0. Thus the value of p is found by solving the equation $0 = 50\,000 \times 1.04^{40} - p \times \frac{1.04^{40} - 1}{0.04}$. Such equations can be readily solved using technology.

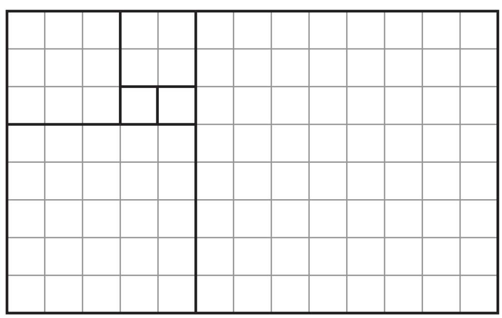
Student activity 5.5

The first few rows of Pascal's triangle, that is, the coefficients of terms in the binomial expansion of $(x + y)^n$, are, in *left-justified* format:

1						
1	1					
1	2	1				
1	3	**3**	1			
1	**4**	6	4	1		
1	5	10	10	5	1	
1	6	15	20	15	6	1

If sums are formed along *diagonals* from *top right* to *bottom left* these generate the values of the Fibonacci sequence {1, 1, 2, 3, 5, 8, 13 …}. The diagonal whose sum generates the term $F(6) = 8$ is shown in bold. As the terms of the rows of Pascal's triangle give values of $\binom{n}{r}$, the following relationship also holds: $F(6) = 8 = \binom{5}{0} + \binom{4}{1} + \binom{3}{2}$.

The following diagram shows a nested series of rectangles with side lengths corresponding to the values of the first n terms of the Fibonacci sequence:



Student activity 5.6

Students should, in the first instance, experiment with values of the parameters within the ranges of those given in the text examples. They can then try other combinations of values.

REFERENCES

Aczel, J & Dhombres, J 1989, *Functional equations in several variables,* Cambridge University Press, Cambridge.

Aczél, J 1966a, *Lectures on functional equations and their applications*, Academic Press, New York.

Aczél, J 1966b, *On applications and theory of functional equations*, Elemente der Mathematik vom Höheren Standpunkt aus, vol. 5, Basel and Stuttgart.

Aigner, M & Zeigler EM 2004, *Proofs from the book* (3rd edn), Springer-Verlag, New York.

Alsina, C 2000, 'Mathematical modelling by means of functional equations: The missing link in the learning of functions' in *Modelling and Mathematics Education ICTMA 9: Applications in Science and Technology,* JF Matos et alia (eds), Horwood Publishing, Chichester.

Binmore, KG 1977, *Mathematical analysis: A straightforward approach*, Cambridge University Press, Melbourne.

Boole, G 1854, *The laws of thought*, Macmillan, London.

Chaitin, G 2000 'A century of controversy over the foundations of mathematics' in *Finite versus Infinite*, CS Calude & G Paun (eds), Springer-Verlag, New York, pp. 75–100.

Cohen, SJ 2002, *Computer algebra and symbolic computation*, Peters, Natick MA.

Crossley, JN 1972, *What is mathematical logic?* Oxford University Press, Oxford.

Eichhorn, W 1978, *Applied mathematics and computation: Functional equations in economics*, Addison-Wesley, Reading MA.

Enderton, HB 1977, *Elements of set theory*, Academic Press, New York.

Estep, D 2002, *Practical analysis in one variable*, Springer-Verlag, New York.

Finney, RL, Thomas, GB, Demana, FD & Waits, BK 1994, *Calculus: Graphical, numerical, algebraic,* Addison-Wesley, New York.

Gleick, J 1987, *Chaos*, Cardinal, London.

Hammel-Garland, T 1987, *Fascinating Fibonaccis: Mystery and magic in numbers,* Dale Seymour, Palo Alto CA.

Hodgson, B & Leigh-Lancaster, D 1990, *Space and number,* Jacaranda Press, Brisbane.

Hoggatt, VE 1969, *Fibonacci and Lucas numbers*, Houghton Mifflin, New York.

Lipschutz, S 1974, *Linear algebra*, McGraw-Hill, New York.

Martin, GE 1982, *Transformation geometry: An introduction to symmetry*, Springer-Verlag, New York.
Mendelson, E 1979, *Introduction to mathematical logic*, 2nd edn, Van Nostrand, New York.
Potts, RB 1985, Discrete Mathematics. Plenary presentation, *Proceedings of the Fifth International Congress on Mathematics Education*, M Carss (ed.), Birkhäuser, Boston.
Skemp, R 1989, *Structured activities for primary mathematics: How to enjoy real mathematics, vols 1&* 2. Routledge, London.
Stephens, M 2004, 'Relational thinking about number as a bridge to algebraic reasoning' in *Making mathematics vital: Proceedings of the 20th biennial Conference of the Australian Association of Mathematics Teachers*, M Coupland, J Anderson & T Spencer (eds), AAMT, Adelaide.
Stillwell, J 1999, *Number*, Springer-Verlag, New York.