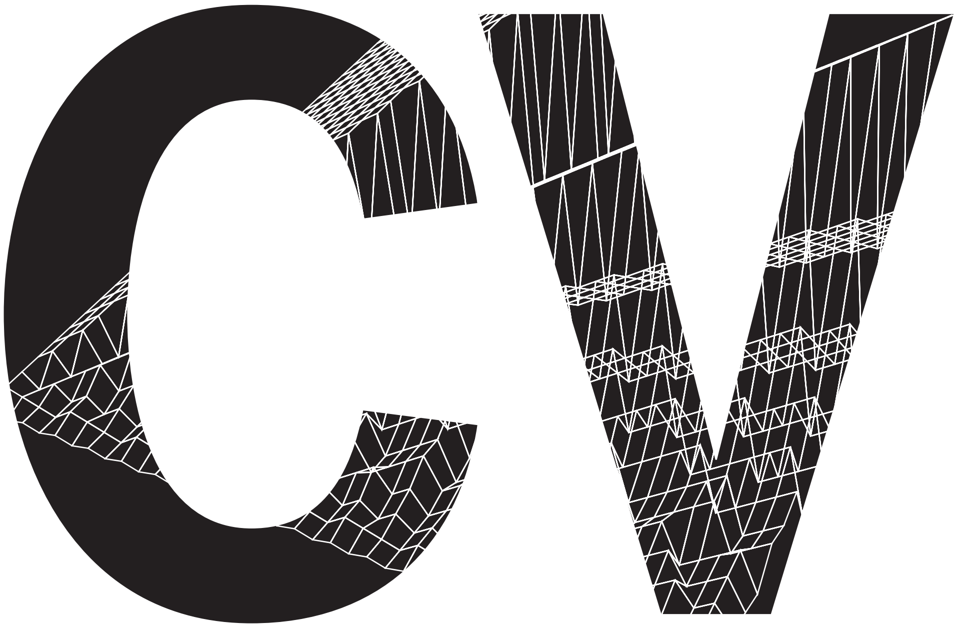
Complex Numbers and Vectors

Les Evans

MathsWorks for Teachers



Complex Numbers and Vectors

Les Evans

MathsWorks for Teachers

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

Edited by Marta Veroni and 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 Pty Ltd

National Library of Australia Cataloguing-in-Publication data:

Evans, Les.
Complex numbers and vectors.

ISBN 0 86431 532 5.

1. Numbers, Complex. 2. Vector analysis. I. Title.
(Series: MathsWorks for teachers).

512.788

Visit our website: www.acerpress.com.au

CONTENTS

This page intentionally left blank

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.

ABOUT THE AUTHOR

Les Evans is an experienced senior secondary mathematics teacher and school director of curriculum, and has been a setter and assessor of formal system extended assessment tasks. His mathematics and education interests are in drawing together different areas of study within problem solving and modelling theoretical investigations and practical applications, and he has extensive professional development experience in these areas.

CHAPTER 1

COMPLEX NUMBERS AND VECTORS IN THE SECONDARY CURRICULUM

INTRODUCTION

Complex numbers and vectors are both important areas of study within the senior secondary mathematics curriculum. They are particularly significant for those students wishing to undertake further study in mathematics or in disciplines that require a strong background in mathematics. These students typically study specialist or advanced mathematics subjects in their senior mathematics curriculum. However, the advent of more sophisticated hand-held technologies over the past decade or so has meant that students from mainstream function, algebra and calculus courses also come across complex numbers as roots to certain types of algebraic equations in the analysis of polynomial functions.

The rationale for the inclusion of *complex numbers* in the curriculum is often related to:

- arguments for completeness of algebraic analysis of polynomial functions and the solution of related equations
- consideration of certain types of transformations of the (complex) plane, in particular those involving combinations of dilations and rotations, as well as some curves and regions in the complex plane

Students typically encounter complex numbers in the guise of some kind of special number that enables one to extend certain algebraic manipulations on quadratic functions of a real variable with real coefficients to ensure that the rule of any quadratic function $q(x) = ax^2 + bx + c$ can be expressed as a product of two linear factors, and the equations $q(x) = 0$ always has two (not necessarily distinct) roots. This is a restricted case of a more general result, indeed, one of the major results of mathematics—the Fundamental Theorem of Algebra. Ironically, the typical school context for the introduction of complex numbers is quite removed from the original historical context—the search for a general formula for the solution of cubic polynomial equations in a single variable. While complex numbers can happily be ignored for a quadratic polynomial function if real roots are required, this is not the case for the roots of a cubic polynomial function.

The rationale for the inclusion of *vectors* in the curriculum is often related to:

- geometry and proof as well as the representation of coordinates and certain types of transformations of the plane (see, for example, Adler 1966; Martin 1982)
- representation and analysis of planes, shapes and curves in three-dimensional space
- applications to the analysis of co-planar forces in statics and dynamics

Vectors are also applied to geometry, kinematics and vector calculus in three dimensions. Indeed, the two areas of study, complex numbers and vectors, overlap with respect to vector representation of the additive structure of complex numbers in the complex plane (complex vectors). These areas of study, along with Boolean algebra in logic, are often used to provide examples of mathematical structures that differ from those underpinning the number systems typically studied in school mathematics. In particular, complex numbers are examples of an *unordered field* in analysis, while vectors are examples of a *vector space* in linear algebra.

Modern technologies, such as graphics or CAS calculators, dynamic geometry systems and computer algebra systems (CAS) provide a range of functionalities for analytical, numerical and graphical computation with complex numbers and vectors. Teachers and students can use these to deal with each of these aspects of the analysis of complex numbers and vectors and related two-dimensional and three-dimensional curves and regions.

This resource provides background information for teachers on both complex numbers and vectors, with consideration of the geometry of the plane as a common theme, and integrates historical material and the use of technology throughout.

COMPLEX NUMBERS

Various different numbers and number systems have been used by humans in different societies and cultures from pre-history to the modern era. Tally sticks such as the thigh bones of animals from Palaeolithic times show arrangements of vertical incisions, often in groups of five, or in what appear to be prime number groupings (see Struick 1948). The relationship between language and words for numbers is fascinating. Linguistic analysis provides clues to how early counting systems evolved, and their descendants can be found in modern languages (see Deakin 1996). That is, they show how quantitative forms evolved in language from earlier qualitative notions, including simple fractions such as one-half, where the representative word seems to have emerged independently from that of the corresponding whole number.

Words for numbers in modern languages that have evolved from Proto-Indo-European roots show that bases 2, 5, 10, 12, 20 and 60 have been part of working with numbers for a long time. As early civilisations emerged and historical records were kept, it becomes clear that fractions and operations with fractions, and, in some cases an explicit concept of zero, were part of these cultures. In the Western tradition of mathematics, the archaeological and historical emergence of number follows a similar but somewhat different order from that of the sequence of development in the contemporary curriculum for the compulsory years of education. It can be argued that a plausible broad 'historical' overview for the development of *number* might be as follows:

- Early ***numberless*** counting (30 000–40 000 BC: pre-history, archaeological artefacts include animal thigh bones with grouped vertical incisions, grouped in fives or less, with up to 20 in a group)
- Counting using ***words*** (?–5000 BC: one, two, three (sometimes four), then many. Multiple words for more than one (duo, brace, yoke, pair) associated with different objects of aggregation (oxen, dogs). Composites using ones and twos. Psychological studies in perception suggest that four objects is the maximum most humans can visually identify at a glance *without* 'counting' and similarly for birds such as crows. Linguistic evidence supports this by analysis of plural forms (special forms of plural for 'two' objects such as identical items 'feet' and of two identical parts 'scissors', when compared with 'many' objects which have a generic plural form) and the construction of words for simple fractions such as $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{5}$ (the names for one-*third* and one-*fifth* reflect their reciprocal relationship with *three* and *five*, while the word 'one-half' or 'halve' does not come from *two*; that is, it likely had its 'own name' before that of the counting sequence beyond three).

- ***Natural*** numbers (several millennia BC: ancient civilisations such as Babylonia, Egypt); various bases used in computations and timekeeping, e.g. 2, 5, 10, 12, 20, 24 and 60
- ***Positive rational*** numbers in fractional form (several millennia BC: Babylonia sexagesimal—base 60—fractions; Egypt using unit fractions)
- ***Irrational*** surds such as $\sqrt{8}$ and irrational numbers such as the golden ratio phi, φ, and the ratio of a circle perimeter to its circumference, pi, π (about 500 BC: ancient Greece)
- ***Integers*** (around 1200 AD: pre-Renaissance Italy)
- ***Complex*** numbers such as $1+\sqrt{-3}$: (early to mid 1500s Renaissance Italy)
- ***Decimal*** arithmetic for ***rational*** numbers (mid 1500s: England)
- ***Infinitesimals*** (mid to late 1600s, although it can be argued that there are earlier antecedents back to the work of Archimedes, the explicit use of infinitesimals—the infinitely small—is found in the work on calculus of Leibniz (dx) and Newton (moments))
- ***Quaternions*** as an extension of complex numbers with applications in physics (mid 19th century, Hamilton)
- The theory of ***real*** numbers (mid to late 19th century, Dedekind)
- ***Transfinite*** cardinal and ordinal numbers in set theory (late 19th century, Cantor)
- ***Non-standard*** natural numbers and real numbers, and ***infinitesimals*** as non-standard real numbers (mid 20th century, Robinson)

There are also many other aspects of interest in the development of number which have varied considerably with respect to their emergence (or not) in different societies and cultures throughout history, such as:

- whether a *place value* system for number is used
- the concept of *zero*
- the notion of *co-measurability* (given any two lengths, does there exist a common unit of which the given lengths are whole number multiples)
- the straight edge (unmarked ruler) and compass *construction* of line segments whose lengths determine the 'location' of certain numbers on a straight line with respect to a fixed origin
- particular *types* and sequences of natural numbers such as odd, even, triangular, square, pentagonal, hexagonal, prime and perfect

In the early years of school mathematics, students typically work through the natural number sequence, with zero making its appearance somewhere along the line. Attention is given to student ability to be able to conceptualise each number as a *reified* construct which can then be further operated within its own right, rather than just the name for the next object in a sequence.

Thus, it is not surprising to find young students who can list the names of numbers in correct sequence using memorisation of linguistic patterns far beyond their ability to 'know' what the number they are *naming* actually 'is', and before any concrete experience with representations of the actual number.

The process of naming involves familiarity with both is description in *words* and its *numeral* designation. Thus, the fourth natural number (if zero is taken as the starting number) is 'three', and its numeral designation might be 三 (Chinese), **III** (Roman) or **3** (modern Hindu–Arabic). Early number work focuses on developing student familiarity and confidence with natural numbers as reified constructs, as well as their names in sequence. On the other hand, the ability to conceptualise the existence of large natural numbers independent of any experience with concrete representation is an essential aspect of number associated with place value representation and knowledge of the natural numbers as a successor sequence starting at zero (or one as may be the case) and increasing by a single unit from one term of the sequence to the next.

The ability to find suitable representation for various conceptualisations of 'number' and work effectively with related computations has been central to the history of human mathematical development. Comprehension of this process, with its triumphs and pitfalls, is also important to students learning about number in mathematics. Whole numbers and simple fractions have natural *models* and interpretations that humans can generally access fairly easily. That is, one can form collections of any size and form various fractions of these, or use lengths, areas and the like to provide models for a unit and fractional parts of it (although there are certainly various subtleties in this process, see Skemp 1987). Irrational numbers such as square roots can be modelled as measuring the length of certain line segments (although the Pythagoreans had metaphysical difficulty with the concept of non-commeasurability); however, it was some time before negative integers were accepted, following their introduction in the 11th century for accounting practices by Leonardo (Fibonacci) of Pisa. René Descartes was not impressed with the notion of negative integers—he only used positive lengths in his formulation of what would later be called the Cartesian coordinate system in his honour.

While complex numbers were first used in the 15th century by Renaissance Italian mathematicians, this was as a device for manipulative convenience—models for complex numbers were not developed until the early to mid 19th century. This was done independently by Gauss, using the *geometry* of the complex plane (1832), and Hamilton, using *ordered pairs* of real numbers (1835). These ideas were further refined with the development of the Argand diagram, following the work of Argand and Wessells in the latter part of the 19th century. The notion of a plane (the *complex* plane) is central to comprehending the various models for complex numbers using algebraic objects, ordered pairs, complex vectors, transformations and matrices.

Indeed, investigations into *arbitrary* sorts of abstract mathematical objects and structures by English and European mathematicians in the early to mid 19th century led fairly directly to the development of both complex numbers and vectors as they are known today. The thinking underpinning this approach is well expressed by the English mathematician and logician George Boole in the introduction to his *Mathematical Analysis of Logic* (1847):

> They who are acquainted with the present state of the theory of Symbolical Algebra, are aware, that the validity of the processes of analysis does not depend on the interpretation of the symbols which are employed, but solely upon the laws of their combination. Every system of interpretation which does not affect the truth of the relations supposed, is equally admissible, and it is thus that the same process may, under one scheme of interpretation, represent the solution of a question on the properties of number, under another, that of a geometrical problem, and under a third, that of a problem of dynamics or optics.

Several of Boole's contempories, such as Augustus De Morgan and William Rowan Hamilton, investigated generalisations of everyday arithmetic—Hamilton, successfully, in 1835 with the *ordered pair* (which he called ordered couples) representation of *complex numbers*, and *quaternions* (1853), while in the early 1850s De Morgan also wrote on complex numbers, which he called the 'double-algebra' as part of his more general investigations of algebraic structure. They both attempted to find, without success, an algebra of 'ordered-triples'. However, Hamilton did develop the theory of *vectors* (the name is due to Hamilton) as part of his study of *quaternions*, as, independently, did the German mathematician Hermann Grassman—vectors were part of his *theory of extensions*. In contemporary mathematics, vectors, along with matrices, are part of what is known as the study of *linear algebra*. Thus, both complex numbers and vectors have substantial common as well as distinctive roots in the history of mathematics and in their applications in the contemporary discipline of mathematics.

Where complex numbers, vectors and possibly also matrices are included as areas of study or topics in the implementation of a senior secondary mathematics curriculum, there is opportunity for material from one or more of these areas of study or topics to be used to support the development of related concepts, skills and processes in the other areas of study or topics. This may occur either within a single course of study which incorporates complex numbers, vectors and matrices, or across two related courses of study, where a course with advanced material builds on a course of core material.

For example, matrices provide a natural representation for two-dimensional coordinate vectors when they are being considered in terms of transformations of the plane via translations. They also have natural applications to the analysis of bases for vector spaces and linear independence. Matrices can also be used to provide a model for complex numbers arising from consideration of certain transformations of the plane. Similarly, two-dimensional coordinate vectors in the complex plane provide a convenient geometric model for the additive structure of complex numbers. In practice, using these connections to advantage requires some careful sequencing of material, and possibly also review of material covered earlier.

It is often the case that while natural models for vectors and matrices, and intuitive interpretations for these in practical contexts are usually accessible to students, this is not so readily the case for complex numbers. Thus a variety of approaches are sometimes employed to help students develop their understanding of important concepts in this area of study. These include:

- developing the structure of complex numbers and their arithmetic as analogous to that of certain surds, for example, consideration of the structure of the set of surds defined by $a + bp$ where a and b are rational numbers and $p^2 = 2$, to explore the definition of complex numbers as the set of numbers defined by $a + bi$ where a and b are real numbers and $i^2 = -1$ with corresponding arithmetic operations
- the use of complex vectors to explain the additive structure of complex numbers, and the interpretation of the multiplicative structure of complex numbers in terms of their polar form and compositions of dilations and rotations in the complex plane
- the use of the abstract ordered pair model for complex numbers, as developed originally by Hamilton
- the use of the sub-ring of the ring of 2×2 matrices with the usual operations of matrix addition and multiplication, where $a + bi$ is represented in terms of the matrix

$$a\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} + b\begin{pmatrix}0 & -1\\1 & 0\end{pmatrix} = \begin{pmatrix}a & -b\\1 & a\end{pmatrix} = aI + bJ$$

where a and b are real numbers. The corresponding structure is isomorphic to the field of complex numbers, and can be interpreted in terms of transformations of the plane.

When one construct, such as a vector or a matrix, is used to model another construct, such as a complex number, care must be taken to ensure that students do not confuse conventions of symbolic representation used for one construct with those used for another. For example, the vector notation $a\mathbf{i} + b\mathbf{j}$, with **bold** or tilde underscore used to indicate a vector quantity, is very similar in appearance to the complex number notation $a + bi$. In the first context **i** represents a *real unit vector* commonly held to designate a horizontal direction, while in the second context i represents an *imaginary unit vector* along the vertical axis of an Argand diagram. If complex impedance and phasor diagrams in electric circuits are used as an *application context* for complex numbers (which links cartesian, polar and exponential forms), the symbol j is used instead of i to represent the imaginary part of a complex number, because i or $i(t)$ is used to represent *current*. Similarly, within the context of analysis of 2×2 matrices the symbol I is often used to designate the identity matrix for multiplication, and the symbol J is sometimes used to designate a rotation of 90 degrees anticlockwise about the origin. Thus, using these conventions, the matrix equivalent for a complex number would be written as $aI + bJ$ in this notation. While there are advantages in teachers availing themselves of multiple representations of mathematical constructs, and doing so by linking across areas of study or topics, this requires care if these connections are not to be a source of consequent confusion to students.

SUMMARY

- Complex numbers are typically included in senior secondary mathematics curricula to provide:
  - completeness of algebraic analysis of polynomial functions and the solution of related equations
  - representation and analysis of transformations of the (complex) plane involving dilations and rotations with respect to the origin
  - the study of curves and regions in the (complex) plane.
- Vectors are typically included in senior secondary mathematics curricula to provide:
  - a natural representation of coordinates and translations
  - an alternative approach to geometric proof

SUMMARY (Cont.)

  - representation and analysis of planes, shapes and curves in three dimensions
  - the study of kinematics, statics and dynamics.
- Complex numbers can be represented by objects of the form $a + bi$, where a and b are real numbers and $i^2 = -1$, ordered pairs with defined relationships, certain 2×2 matrices, and transformations (rotations and dilations) of the complex plane. These representations are interconnected.
- Complex numbers provide a model of an unordered field (complex algebra).
- Vectors provide a model of a vector space (linear algebra).
- Vectors can be represented by arrows, ordered pairs, ordered triples and their natural extensions, and matrices of order $1 \times n$ (a row vector) or matrices of order $n \times 1$ (a column vector), where n is natural number greater than 1. These representations are interconnected.
- Vectors can be used to model the additive structure of complex numbers, using what are called complex vectors (the ordered pair representation of a complex number in the complex plane interpreted as a vector in this plane).
- Modern dynamic geometry and computer algebra technology can be used to effectively support representation and computation with complex numbers and vectors.

References

Adler, I 1966, *A new look at geometry*, John Day, New York.
Boole, G 1847, *The mathematical analysis of logic*, Macmillan, London.
Deakin, MAB 1996, 'The origins of our number words', *Australian Mathematical Society Gazette*, 23, 2, 50–66.
Martin, GE 1982, *Transformation geometry: An introduction to symmetry*, Springer-Verlag, New York.
Skemp, R 1989, *Structured activity for primary mathematics: How to enjoy real mathematics*, vols 1 & 2, Routledge, London.
Struik, DJ 1948, *A concise history of mathematics*, Dover, New York.

CHAPTER 2

A TALE OF INTRIGUE AND IMAGINATION

THINKING OUTSIDE THE SQUARE

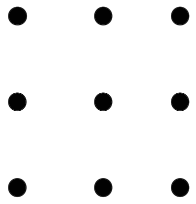
Mathematicians are problem solvers. They pose problems and seek solutions. Good mathematical problems will intrigue and bring our imaginations into play. When solving problems we will sometimes experience the tension of disappointment or, hopefully more often, the joy of discovery.

The best problems can appear deceptively simple, but can create the need to explore new ways of thinking about our world and how we choose to describe it. They push at the boundaries of our knowledge and ask us to use our imaginations to create that which has never existed. We can be the creators of new worlds.

It is our imaginations that allow us to break away from the boundaries we set ourselves or those set by others. Have you ever been asked to 'think outside the square'? Have you ever wondered what this demands of us? A small problem may allow an insight. Student Activity 2.1 can be used to illustrate this point to students.

STUDENT ACTIVITY 2.1

Using only four lines, and without retracing a path or lifting your pencil from the page join the nine dots in the diagram below.



To solve this problem we need to use our imagination and think of solutions that take us beyond the boundaries that seem to have be created by the nine dots. On too many occasions it is the boundaries we have created for ourselves that constrain our world. Thinking outside the square demands that we recognise the boundaries we set for ourselves and have the courage to move beyond them. Too many of the boundaries that confine our thinking are of our own making.

To have the courage to think outside the square, we need to be intrigued by a problem. This intrigue will encourage us to use our imaginations to find solutions which are beyond our current view of the world. This was the challenge that faced mathematicians as they searched for a solution to the problem of finding meaning for the square root of a *negative* number, in particular $\sqrt{-1}$.

Before we can find the solution to the square root of a negative number, we need to understand some of the constraints that hampered its discovery. We should begin by travelling back in time to a religious sect that helped create modern analytical mathematics: the Pythagorean School. It was this sect that set the path that allows us to understand the square root of *positive* numbers. While such a concept seems almost trivial to we students of mathematics in the 21st century, two millennia ago it was a tale of intrigue and imagination.

ALL IS NUMBER

'All is number' was the motto of the Pythagorean School. Number was believed to have mystical powers and the Pythagoreans based their philosophy and way of life on it (Kline 1972). All numbers held special significance (see Table 2.1 for some examples).

When considering and exploring the square roots of positive numbers, the Pythagoreans' mystic belief in number acted as a boundary we should explore if we are to understand the intrigue of the square root of a negative number. Moving beyond the boundary they had set themselves took courage. It was a life-threatening situation. The intrigue began with the simple problem of expressing $\sqrt{2}$ as a number as the Pythagoreans understood them.

In Pythagoras' time the term 'number' was used only for the positive integers. *Integer* is a Latin term that means 'whole, untouched, and unharmed'. Fractions, or the divisions between integers, were not considered to be separate from them, but a ratio or relationship between two integers. Thus $\frac{1}{2}$ is the ratio of the two integers, 1 and 2, and can be expressed as 1:1.

Table 2.1: Some significant numbers (Boyer 1985)

One	It is the generator of all numbers, and thus the number of reason.
Two	All even numbers were considered to be female, so 2 was the first female number. It was the number of diversity and opinion.
Three	All odd numbers were considered to be male. Three was the first male number. It was the number of harmony because it combined reason (1) and diversity (2).
Four	It was the number of justice and retribution. It is the first square number (22), so we have a squaring of accounts.
Five	The number of marriage, it is created by combining the first true male number (3) with the first true female number (2).
Six	The number of creation, it is marriage (5) combined with the generator (1).

Under this constraint their motto 'all is number' becomes more significant. It means that all numbers including the gaps between whole integers are either a simple integer or a ratio of integers (fractions). They believed that $\sqrt{2}$ would also fall into this pattern.

It is not hard to understand why they believed this to be true. Consider for a moment a number line and in particular a section of it between zero and 1. When you start to allocate fractions to this smaller section (Figure 2.1), it soon becomes apparent that there are an infinite number of fractions between zero and 1, for example the sequence $\left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5} \ldots\right\}$. There are many other sequences similar to this one, so it would be hard to imagine that there would be room for any other numbers.

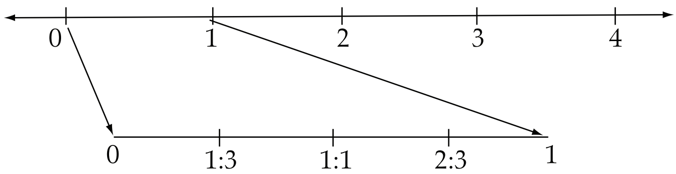


Figure 2.1: The allocation of fractions (as ratios) to the number line

The basic belief of all Pythagoreans was that the essence of all things is explainable by the intrinsic properties of whole numbers or their expression as ratios. Can you imagine the intrigue that would surround the virtual destruction of the foundation of the Pythagoreans' faith in integers and their ratios? This happened when one of their own suggested that there are numbers that cannot be expressed as a ratio of whole numbers. One such number is $\sqrt{2}$.

It is difficult to comprehend the tension of this discovery and the efforts of the Pythagorean School to hold it as their little secret. It is rumoured that any Pythagorean who released this information to an non-Pythagorean was punished by drowning.

INTRIGUE AND $\sqrt{2}$

There are two challenges when we consider the intrigue that surrounded $\sqrt{2}$. The first is to locate it on the number line. The second is to prove that it cannot be represented as a ratio of integers.

We will begin by finding the location of $\sqrt{2}$ on a number line. To do this we will need to explore a geometric representation of $\sqrt{2}$. Square roots can be geometrically represented as the length of a line segment—in particular, as the length of the side of a square that has an area equal to the square of the root under consideration (Figure 2.2). However, this does not give us the exact length of $\sqrt{2}$, only an indication that it may be possible.

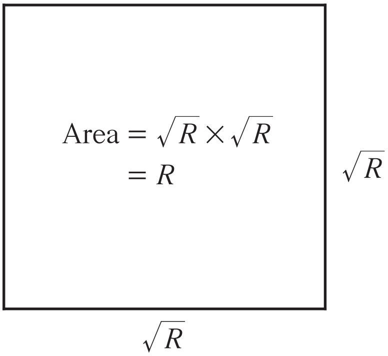


Figure 2.2: A geometric representation of the square root of a number

Socrates (Plato, 1871/Meno) gave us a hint on how to find the exact length of $\sqrt{2}$ when he asked the slave of one of his friends to construct a square with double the area of a given square. After some support from Socrates, the slave presented a solution which, in modern terms, is shown in Figure 2.3.

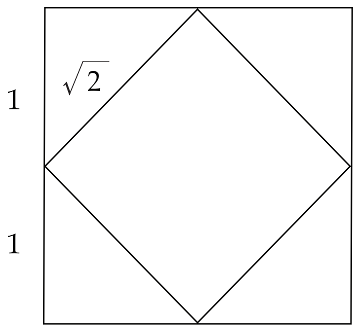


Figure 2.3: Construction of a square that is double the area of a given square

We can use Figure 2.3 to place $\sqrt{2}$ on the number line, using a compass and straight edge construction, as shown in Figure 2.4.

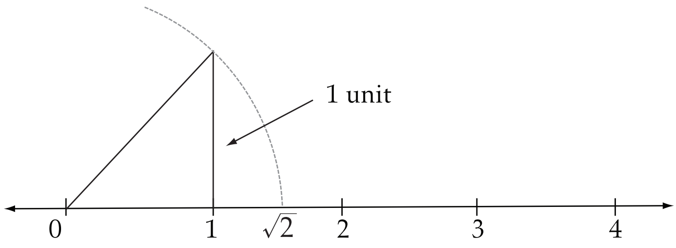


Figure 2.4: Placing $\sqrt{2}$ on a number line

It is important that students become familiar with the actual placement of such numbers on a number line using geomety. Student Activity 2.2 allows students an opportunity to practise this skill.

STUDENT ACTIVITY 2.2

Locate the following numbers on a number line.

a $\sqrt{3}$ b $\sqrt{5}$ c $\sqrt{6}$

d $\sqrt{7}$ e $\sqrt{8}$

Now that we have located $\sqrt{2}$ on the number line, if we can express it as a ratio of whole numbers then the Pythagorean motto 'all is number' will stand. However, if we can show that it cannot be expressed as a ratio of integers, the

Pythagoreans will need to join us and accept the challenge of looking beyond the borders we set for our imaginations.

The proof that $\sqrt{2}$ is not rational is relatively easy if we use modern analytical techniques. While the mathematics is easy, some students may find the ideas behind the proof challenging. It is a proof by contradiction, which relies on assuming that something is true and showing that this assumption is not true. Thus, it is not a constructive proof since we still do not 'have' $\sqrt{2}$ at the end. While this may be a challenge to some students, the proof was more difficult using the geometric approach required in the times of Socrates and Pythagoras.

We should start with a square (Figure 2.5) with lengths as shown.

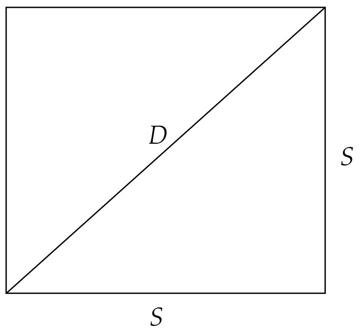


Figure 2.5: A square of dimension *S* and diagonal *D*

Let us assume that $\frac{D}{S} = \frac{p}{q}$, where p and q are positive integers with no common factor.

Using Pythagoras' theorem:

$$\begin{aligned} D^2 &= S^2 + S^2 \\ 2 &= \frac{D^2}{S^2} \\ &= \frac{p^2}{q^2} \\ \therefore p^2 &= 2q^2 \qquad (1) \end{aligned}$$

This suggests that p^2 must be even, hence p must be even.

Student Activity 2.3 could be used to encourage students to carry out simple proofs and thus develop an understanding of the notion of generalisation and proof.

STUDENT ACTIVITY 2.3

a Show that for all even numbers, n, n^2 is also even.
b Show that for all odd numbers, n, n^2 is also odd.
c Show that when two even numbers are multiplied the result will be even.
d Show that when two odd numbers are multiplied the result will be odd.
e Show that the product of an odd and even number is even.
f What will be the result when an odd and an even number are added together?
g What will be the result when two odd numbers are added together?
h What will be the result when two even numbers are added together?

If p is even, then q must be odd, otherwise p and q would have a common factor of 2.

Let $p = 2r$, and substitute into equation (1).

We gain $4r^2 = 2q^2$

$$\Rightarrow q^2 = 2r^2$$

This suggests that q^2 must be even, hence q must be even.

However, we have previously shown that q must be odd. We know that an integer cannot be both odd and even. This means that the assumption that $\sqrt{2}$ can be express as a ratio of whole numbers must be false. The nature of this proof has caused concern for some mathematicians. It is not a positive proof; rather, it leads to a contradiction that allows us to draw a favourable conclusion. Fortunately it is possible to create a geometric construction that models $\sqrt{2}$.

This proof together with the ability to place $\sqrt{2}$ on the number line was a blow to the Pythagoreans. It is little wonder they wanted to keep it as their little secret.

The simple act of proving that $\sqrt{2}$ cannot be expressed as a ratio of whole numbers pushed at the boundaries of mathematics at the time. In a time where such discoveries were accompanied by large risk, it took courage, intrigue and imagination to create another world of mathematics. When we want to describe this new world, we may need to seek out new mathematical tools. The mathematics of investigation and proof during the time of the Pythagorean School relied on geometry. It allowed them to visualise the square roots of positive numbers using squares, but it had limitations which created boundaries to other discoveries.

The geometry that allowed the recognition and placement of the square roots of numbers that cannot be expressed as a ratio of whole numbers acted as a boundary to the recognition that we may also be able to find the square root of a negative number.

We can use geometry to show that $(a + b)^2 = a^2 + 2ab + b^2$ (Figure 2.6).

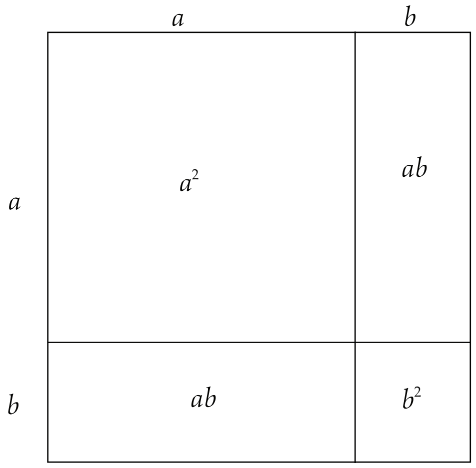


Figure 2.6: Geometric representation of $(a + b)^2$

In a similar way we can use geometry to show that $(a + b)(a - b) = a^2 - b^2$ (Figure 2.7).

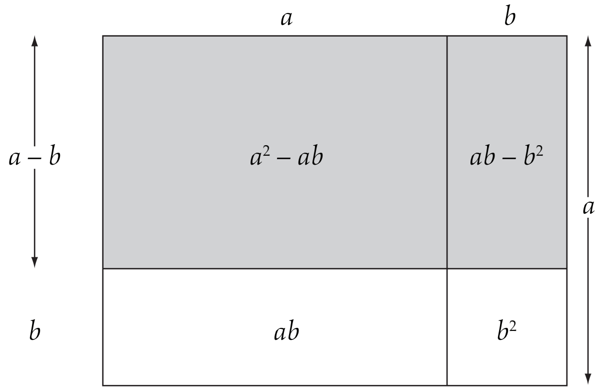


Figure 2.7: Geometric representation of $(a + b)(a - b)$

Geometry offers a powerful insight into the factorisation of quadratic expressions. It creates opportunities to understand the process by using a concrete model. Student Activity 2.4 is included to allow students an appreciation of the use of geometry to represent the factorisation of quadratic equations. It also offers an insight into the challenge of considering the square roots of negative numbers, when the problem is considered from a geometric perspective.

STUDENT ACTIVITY 2.4

Use geometry to show that:

a $a(a+b) = a^2 + ab$

b $a(a-b) = a^2 - ab$

c $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$

d $a^3 - b^3 = (a-b)(a^2 + ab + b^2)$

The beauty of geometric proofs is that we can visualise abstract concepts as lengths, areas and volumes that can be manipulated. Geometry allows us to work with a number such as $\sqrt{2}$ and realise that it has an exact length, although we would not be able to measure it exactly using fractions or decimals.

It is an interesting notion that we can cut a length of rod that is exactly $\sqrt{2}$ in length, but we can only approximately measure its length if we were to use a metric rule. This would still be true if we decided to work to nanometres or even smaller units. It is little wonder that the Pythagoreans found it difficult to imagine such a number.

Now that we have worked with and developed an understanding of the square root of a positive number, we need to be bold and take another step. We need to consider the square root of a negative number.

Unfortunately, if we were to use geometry to explore the concept of the square root of a negative number, we would be setting a boundary to our imagination that would be difficult to cross. To represent $\sqrt{-1}$ using geometry would require us to draw a square with each side length being less than zero. To be asked to draw a square with side length less than zero sounds similar to the Zen Buddhists asking ‘What is the sound of one hand clapping?’

SUMMARY

- There are many boundaries to the creation of new mathematics. Some are our own creation, related to our beliefs, values and preferences, while others relate to limitations of mathematical tools that are available.
- For the Pythagoreans, boundaries existed because of their definition of number and the mystical powers they attributed to whole number.

SUMMARY (Cont.)

- The proof that $\sqrt{2}$ could not be represented as a ratio of whole numbers expanded the boundaries of their mathematics. Encouraging students to work through this proof can create the opportunity for students to improve their understanding of mathematical generalisation and proof.
- Geometry can be used to represent the expansion of quadratic and cubic polynomial expressions, this approach may enable students to develop their understanding of the processes of factorisation and expansion.

References

Boyer, CB 1989, *A history of mathematics*, 2nd edn, UC Merzback (revn ed.), Wiley, New York.
Kline, M 1972, *Mathematical thought from ancient to modern times*, Oxford University Press, New York.
Plato, 1871, *The Meno*, B Jowett (trans.), web edition at http://www.mdx.ac.uk/www/study/xplameno.htm

Further reading

Bell, ET 1937, *Men of mathematics*, Simon & Schuster, New York.
Dedron, P & Itard, J 1973, *Mathematics and mathematicians*, vols 1 & 2, JV Field (trans.), Open University Press, Milton Keynes.

Websites

http://www.math.mcgill.ca/labute/courses/133f03/VectorHistory.html
This website contains an overview of the history of vectors.
http://www-groups.dcs.st-and.ac.uk/~history/HistTopics/Abstract_linear_spaces.html
This website contains an overview of the history of abstract linear spaces.
http://history.hyperjeff.net/hypercomplex.html
This site presents a timeline of the developments in complex numbers.

CHAPTER 3

SECRECY, CONTRIVANCE AND INSPIRATION

RENAISSANCE ITALY AND MATHEMATICAL DUELS

The historical development of complex numbers owes a great deal to contests between mathematicians as they attempted to prove their skills by solving cubic polynomial functions. It is worth exploring this contest with students, using it as stimulus material for the further exploration of complex numbers.

It is appropriate to begin with two Italian mathematicians, Cardano (1501–76), the illegitimate son of a Milan lawyer, and Fontano, (1499–1557), the son of a poor postman from Brescia, who acted together and separately to draw complex numbers into the world of mathematics. It was through their efforts that $\sqrt{-1}$ became easier to imagine, and started to take form.

Their achievements were clouded by jealousy, betrayal, resentment, hatred and overconfidence. It is hard to imagine that the challenge of solving cubic polynomial equations could create a level of animosity that would cause Fontano to publish a work that contained personal malicious insults directed at Cardano. Yet the competition that drove both men led to the formulation of the algebra of complex numbers.

Fontano, the son of a postman, faced death when only 12 years old. Brescia was captured by the French and the town was put to the sword. A French soldier thrust a his sword into the young Fontano's face, breaking his jaw and destroying his palate. He survived through the love and care of his mother who nursed him back to health. Unfortunately, he was left with a speech impediment and a ugly scar across his face. It is sad to realise that his nickname Tartaglia, meaning stammerer, is the name by which he is remembered.

Fontano taught himself mathematics, and decided to use his ability to achieve fame and fortune. He achieved fame by becoming part of the great mathematical challenge of the early 16th century, solving cubic polynomial equations. Mathematicians of the time took delight in challenging each other to solve sets of equations. The rule of the game was to find the roots of a cubic polynomial equation by forming an expression in terms of the coefficients using the four operations of arithmetic and the roots of integers.

The greatest conquest of Fontano was the defeat of Fior, who was taught a technique of solving cubic polynomial equations by del Ferro. Del Ferro was believed to be the first mathematician to discover a general method for solving cubic polynomial equations. He was highly secretative about the process, and only passed on his method to his student Fior when on his deathbed.

Fior was extremely aware of the secretative nature of his teacher, so was supremely confident that he held an advantage in any contest involving the solving of polynomial cubic equations. He became cocky, and informed all who would listen of his skill and ability. This encouraged the mathematical community to arrange a challenge for Fior. Fontano was to be their champion. Fior and Fontano each composed 30 equations for the other to solve within 50 days.

Fior was overconfident and boasted that Fontano would never be able to solve the equations he had set him. Two hours later Fontano had solved his 30 equations. His success made him famous, but fortune still eluded him.

Fontano's newly found fame reached the ears of Cardano, who was considered to be the most able mathematician in Italy—and he knew it. Cardano was outspoken and highly critical of his colleagues and had a habit of making enemies.

Cardano was desperate to know the methods used by Fontano to solve cubic polynomial equations. Fontano did not trust Cardano and intially refused to share his knowledge with him. Cardano used Fontano's desire for fortune to achieve his aim. He suggested that he would use his influence in Milan to gain Fontano a position in the Milan court. Reluctantly Fontano passed his methods on to Cardano after having received a promise that he would never reveal his methods to anyone. Cardano made a very strong commitment for his day:

> I swear to you, by God's holy Gospels, and as a true man of honour, not only never to publish your discoveries, if you teach me them, but I also promise you, and I pledge my faith as a true Christian, to note them down in code, so that after my death no one will be able to understand them.

In 1545 Cardano published his most famous work *Ars Magna*. In this book he broke his promise to Fontano and published his method for solving cubic polynomial equations. To give Cardano some credit, he did acknowledge the process to Fontano.

It is worth visiting Cardano's method for solving cubic equations of the form $x^3 + mx = n$. This is an involved process, so teachers will need to work carefully through the following steps with students.

Let's begin by expanding $(a - b)^3$:

$$(a - b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$$

Hence $$(a - b)^3 + 3ab(a - b) = a^3 - b^3$$

Notice that $a - b$ is a solution to the original equation $x^3 + mx = n$ if:

$$m = 3ab \qquad (1)$$

$$n = a^3 - b^3 \qquad (2)$$

From equation (1) we obtain $b = \frac{m}{3a}$ (3)

Substituting (3) into (2): $n = a^3 - \left(\frac{m}{3a}\right)^3$

$$\Rightarrow a^6 - na^3 - \frac{m^3}{27} = 0$$

If we let $a^3 = t$ we obtain $t^2 - nt - \frac{m^3}{27} = 0$.

This is a quadratic equation that can be solved using the quadratic formula. Thus we can find a by taking the cube roots of the solutions we gain using the quadratic formula. We can use a similar method to find b, and can use both results ($x = a - b$) to find solutions for a cubic polynomial equation.

The cubic equation of the form $x^3 + mx = n$ is known as the *depressed* cubic equation. By the use of a simple substitution the general cubic equation can be expressed in depressed form:

$$0 = ay^3 + by^2 + cy + d$$

By substituting $y = x - \frac{b}{3a}$ this equation becomes:

$$0 = a\left(x - \frac{b}{3a}\right)^3 + b\left(x - \frac{b}{3a}\right)^2 + c\left(x - \frac{b}{3a}\right) + d$$

By expanding and simplifying this equation becomes:

$$0 = ax^3 + \left(c - \frac{b^2}{2a}\right)x + \left(d + \frac{2b^3}{27a^2} - \frac{bc}{3a}\right)$$

This suggests that:

$$m = c - \frac{b^2}{2a} \text{ and}$$
$$n = -\left(d + \frac{2b^3}{27a^2} - \frac{bc}{3a}\right)$$

Student Activity 3.1 provides students with an opportunity to practise these transformations and solutions for specific cases. It would be helpful to students for the teacher to work through a specific example in parallel with the general argument provided earlier.

STUDENT ACTIVITY 3.1

1 Express the following cubic expressions in depressed form.

a $x^3 - 15x^2 + 81x - 175$

b $x^3 + 8x^2 + 25x + 26$

c $2x^3 - 21x^2 + 68x - 29$

d $2x^3 - 25x^2 + 102x - 130$

2 Using the depressed form, find the roots of these cubic expressions.

It is also possible to use computer algebra to solve cubic equations for the general case and for specific examples. It would be opportune to compare results using the methods available to Fontano with those that can be achieved using modern technology. We can use computer algebra systems to find the depressed form of the general cubic equation.

We can also solve cubic equations directly:

```
Solve[ax³ + bx² + cx + d == 0, x]
```

$$\{\{x \to -\frac{b}{3a} - \frac{2^{1/3}(-b^2+3ac)}{3a\left(-2b^3+9abc-27a^2d+\sqrt{4(-b^2+3ac)^3+(-2b^3+9abc-27a^2d)^2}\right)^{1/3}} +$$
$$\frac{\left(-2b^2+9abc-27a^2d+\sqrt{4(-b^2+3ac)^2+(-2b^3+9abc-27a^2d)^{1/3}}\right)}{3\,2^{1/3}a}\},\{x \to$$
$$-\frac{b}{3a} + \frac{(1+i\sqrt{3})(-b^2+3ac)}{3\,2^{2/3}a\left(-2b^3+9abc-27a^2d+\sqrt{4(-b^2+3ac)^3+(-2b^3+9abc-27a^2d)^2}\right)^{1/3}} -$$
$$\frac{1}{6\,2^{1/3}a}(1-i\sqrt{3})\left(-2b^3+9abc-27a^2d+\sqrt{4(-b^2+3ac)^3+(-2b^3+9abc-27a^2d)^2}\right)^{1/3})\},\{x \to$$
$$-\frac{b}{3a} + \frac{(1-i\sqrt{3})(-b^2+3ac)}{3\,2^{2/3}a\left(-2b^3+9abc-27a^2d+\sqrt{4(-b^2+3ac)^3+(-2b^3+9abc-27a^2d)^2}\right)^{1/3}} -$$
$$\frac{1}{6\,2^{1/3}a}(1+i\sqrt{3})\left(-2b^3+9abc-27a^2d+\sqrt{4(-b^2+3ac)^3+(-2b^3+9abc-27a^2d)^2}\right)^{1/3})\}\}$$

It was while solving cubic equations in the depressed cubic form that Cardano realised that particular cubic equations generated unusual solutions that he did not fully understand.

One such result arises when we attempt to solve $x^3 = 15x + 4$.

Using Cardano's method:

$$x^3 - 15x = 4$$
$$m = -15 \text{ and } n = 4$$

Substituting these values into $t^2 - nt - \frac{m^3}{27} = 0$, we gain:

$$t^2 - 4t - \frac{(-15)^3}{27} = 0$$
$$t^2 - 4t + 125 = 0$$

Using the quadratic formula:

$$\begin{aligned} t &= \frac{4 \pm \sqrt{16 - 500}}{2} \\ &= \frac{4 \pm 2\sqrt{4 - 125}}{2} \\ &= \frac{4 \pm 2\sqrt{-121}}{2} \\ &= 2 \pm \sqrt{-121} \\ &= 2 \pm 11\sqrt{-1} \end{aligned}$$

Cardano published a similar result in *Ars Magna*. However, he did not have the benefit of modern notation, which makes his achievement more remarkable. While he did not fully understand complex numbers, he presented the first calculation using complex numbers. To quote him directly:

> Dismissing mental tortures, and multiplying $5 + \sqrt{-15}$ by $5 - \sqrt{-15}$, we obtain $25 - (-15)$. Therefore the product is 40 … and thus far does arithmetical subtlety go, of which this, the extreme, is, as I have said, so subtle that it is useless.

From his comments, it appears that he did not see a place for complex numbers. For this solution to make any sense, we need to search beyond his view of mathematics. Cardano's first step was to *accept that* $\sqrt{-1}$ *existed* and that it *can be used to generate solutions* to cubic polymonial equations. We also must be aware that we cannot obtain this result using real numbers. For this to be true, we must first be comfortable with negative numbers and the notion that $-1 \times -1 = 1$. This is an idea that requires further investigation.

TWO NEGATIVES MAKE A POSITIVE ($-1 \times -1 = 1$)

I can still remember being told, early in my mathematics education, 'all you need to remember is that when you multiply integers, unlike signs give a negative and like signs give a positive'. Does this mean that we should accept without question that $-a \times -b$ will produce a positive result? But all those years ago, it was accepted by the Pythagoreans that $\sqrt{2}$ can be represented as a ratio of whole numbers. Is it too much to ask for a little evidence that $-a \times -b$ does in fact produce a positive result?

Barry Mazur (2003) in his book *Imagining Numbers* offers an interesting proof that $-1 \times -1 = 1$, which is worth considering here. For this proof we need to look to the world of finance.

In the world of finance, if I owe money to someone, from my point of view, this is a negative amount. However, if I am owed money, or have money in my pocket, then this is a positive amount.

From this perspective, it is easy enough to see that if I owe each of my classmates (C) a dollar, then my total debt would be $-1 \times C$ dollars.

What happens if on the way to school I find a dollar and decide to use this dollar to pay off the debt I owe to one of my classmates? I have reduced my total debt by one dollar. This can be expressed in two ways mathematically:

Total debt is reduced by a dollar: $-1 \times C + 1$

Number of debtors is reduced by one: $-1 \times (C - 1)$

Both the above statements are saying the same thing, so we can equate them:

$$-1 \times C + 1 = -1 \times (C - 1)$$

Now there is a simple twist to the story. What if, from the very beginning, I had no classmates to whom I owed a dollar; that is, if the number of debtors C is equal to zero? This would change the above equation to:

$$-1 \times 0 + 1 = -1 \times (0 - 1)$$
$$\Rightarrow 1 = -1 \times -1$$

This means that the dollar I found on the way to school was not owed to one of my classmates, so it can stay in my pocket.

It could be argued that this is all a bit tricky, and has a similar feel to those double negatives in grammar that can cause some of us to be corrected.

How many times have we heard the following exchange?

Parent: We need to talk about …

Child: I didn't do nothing wrong.

Parent: So what did you do?
Child: Nothing?!

Let us look in greater detail at the sentence 'I didn't do nothing wrong' which could be extended to 'I did not do nothing wrong'. The not is a negative that reverses the meaning of a sentence. Our intention is clear when we say, 'I will not go in the water.' It is just as clear when you say, 'I did not do it (the wrong thing).' But isn't this sentence the same as saying, 'I did nothing wrong'? So adding the 'not' reverses the meaning of the sentence. This means that *I didn't do nothing wrong* is equivalent to saying *I did something*. With this in mind it makes sense to ask, *So what did you do?*

In a similar way the action of not (negative) owing (negative) a dollar to a classmate means I get to keep the dollar. So once again a double negative gives a positive result. This result can be proved formally from the axioms of the real number field (see Allendoerfer & Oakley 1963).

This result makes the search for the square root of a negative number more of an intrigue. The implication of $(-\sqrt{2})^2 = (\sqrt{2})^2 = 2$ is that we need to search *beyond* our current view of the world of mathematics to find the square root of a negative number. The answer must be *beyond* the boundaries set by the real numbers.

Even with all the human frailty exhibited by Carbano and Fontano, they should inspire us to search for a possible location of the square root of a negative number. Their work tells us that if we are to grasp this concept we need to take a closer look at our attempts to solve quadratic equations. We are fortunate because we can use René Descartes' great discovery, the cartesian plane, to aid our exploration.

QUADRATIC EQUATIONS AND THE CARTESIAN PLANE

The following discussion is important as a lead-in to the approach normally taken in senior secondary mathematical curriculum, which introduces complex numbers for algebraic completeness of the solutions of quadratic equations.

Many of us will be familiar with the rule of the quadratic function

$$y = ax^2 + bx + c, \text{ where } a \neq 0$$

This can be rewritten as $y = a(x - h)^2 + k$, where (h, k) is the turning point of the parabola that is the graph of $y = ax^2 + bx + c$ (Figure 3.1).

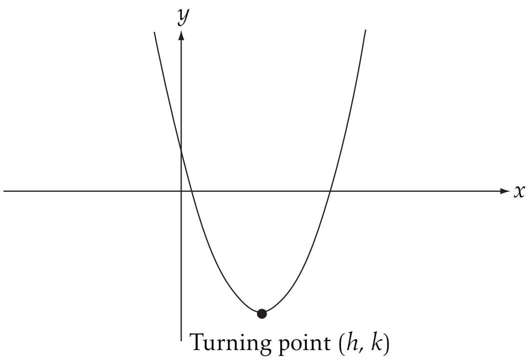


Figure 3.1: Graph of a quadratic with two points of intersection with the *x*-axis

Important aspects of this graph are the points at which it crosses the x-axis—its x-intercepts. We know that the x-axis is the line on the cartesian plane where $y = 0$. Thus, when we want to find the possible x-intercepts we are seeking values for x that allow y to be equal to zero. That is $y = a(x - h)^2 + k = 0$.

Before we attempt to solve this equation, we should consider how values of k might affect our results. There are three possible sets of values for k: $k > 0$, $k = 0$ and $k < 0$.

When $k > 0$, the image of the graph will be shifted so that the turning point is above the x-axis (Figure 3.2).

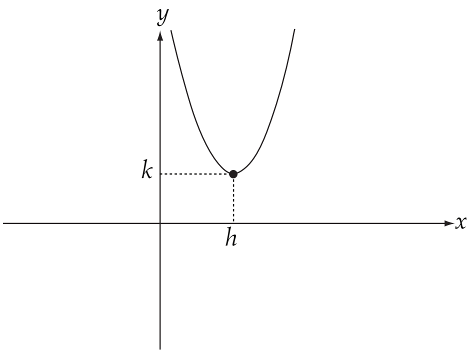


Figure 3.2: Graph of a quadratic where $k > 0$ and $h > 0$

When the parabola is in this position, it does not cross the x-axis, and we would expect that there would be no solutions for $y = ax^2 + bx + c = 0$.

Another position for the parabola occurs when $k = 0$ (Figure 3.3).

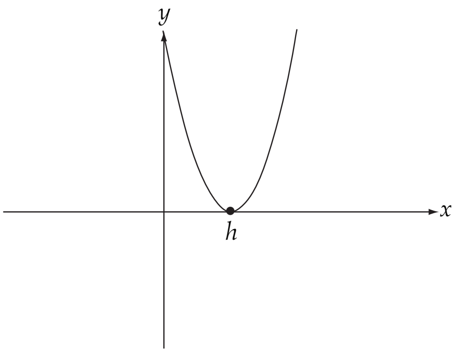


Figure 3.3: Graph of a quadratic where $k = 0$ and $h > 0$

In this case we would expect to have only one solution to the equation $y = ax^2 + bx + c = 0$.

$$\begin{aligned} 0 &= ax^2 + bx + c \\ &= (x - h)^2 \\ \Rightarrow x &= h \end{aligned}$$

The final possibility occurs when k has a negative value (Figure 3.4).

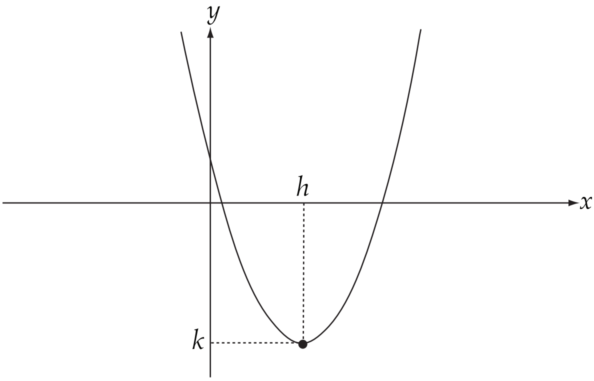


Figure 3.4: Graph of a quadratic where $k < 0$ and $h > 0$

In this case we would expect two solutions to the equation $y = ax^2 + bx + c = 0$:

$$\begin{aligned} 0 &= ax^2 + bx + c \\ &= (x-h)^2 - k \\ &= (x-h)^2 - (\sqrt{k})^2 \\ &= (x - h - \sqrt{k})(x - h + \sqrt{k}) \\ x &= h + \sqrt{k},\ h - \sqrt{k} \end{aligned}$$

Teachers and students familiar with the quadratic formula would realise that $h = \frac{-b}{2a}$ and $k = b^2 - 4ac$. The value of k allows us to discriminate between quadratic equations that will have none, one or two solutions. It is often referred to as the discriminant, Δ, where $\Delta = b^2 - 4ac$. The discriminant can have one of three sets of values.

- It can be greater than zero, $\Delta > 0$, which suggests that a quadratic equation would have two real points of intersection with the x-axis.
- The discriminant can also be equal to zero, $\Delta = 0$, which indicates that the parabola will only intersect with the x-axis at one real point.
- The discriminant can also have a value less than zero, $\Delta < 0$. In this case we would not expect the parabola to intersect with the x-axis, because when $\Delta < 0$ the only solutions that are possible require us to find the square root of a negative number.

We should go into a little more detail. The x-axis is a number line that contains only real number values. We now know that the only solution possible when the discriminant is less than zero would be a number that cannot exist as a *real number*.

Before we abandon this range of values for the discriminant, suggesting that they only provide false answers, we should look again at the problem we began to investigate. The original question was to find values for which $y = 0$, that is, to find the intersection points on the x-axis.

Has the form of the question acted as a boundary to our imaginations? Do we need to stop and consider the possibilities of answers which are beyond those constrained by the real numbers? Like the Pythagoreans, we may be limiting the possibilities because of the way in which we view the world of mathematics. What if the question simply said, find the *factors* of $ax^2 + bx + c$? As an example, what are the *factors* of $x^2 + 2x + 4$?

Using the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

and carrying out the usual computation gives the following:

For $x^2 + 2x + 4$ $\quad a = 1$, $b = 2$ and $c = 4$

$$\begin{aligned} x &= \frac{-2 \pm \sqrt{2^2 - 4 \times 1 \times 4}}{2 \times 1} \\ &= \frac{-2 \pm \sqrt{4 - 16}}{2} \\ &= \frac{-2 \pm \sqrt{-12}}{2} \\ &= \frac{-2 \pm 2\sqrt{-3}}{2} \\ &= -1 \pm \sqrt{-3} \end{aligned}$$

Given that the solutions of an equation formed by equating the rule of a quadratic function to zero are of the form $a \pm b\sqrt{c}$, where a, b and c are real numbers, has factors $(x - a - b\sqrt{c})(x - a + b\sqrt{c})$, this would suggest that we could rewrite $x^2 + 2x + 4$ as $(x + 1 - \sqrt{-3})(x + 1 + \sqrt{-3})$.

This can be checked for consistency by expanding $(x + 1 + \sqrt{-3})(x + 1 - \sqrt{-3})$ using the distributive property:

$$\begin{aligned} (x + 1 + \sqrt{-3})(x + 1 - \sqrt{-3}) &= x(x + 1 - \sqrt{-3}) + 1(x + 1 - \sqrt{-3}) + \\ &\quad \sqrt{-3}(x + 1 - \sqrt{-3}) \\ &= x^2 + x - x\sqrt{-3} + x + 1 - \sqrt{-3} + \\ &\quad x\sqrt{-3} + \sqrt{-3} - (\sqrt{-3})^2 \\ (x + 1 + \sqrt{-3})(x + 1 - \sqrt{-3}) &= x^2 + 2x + 0 + 1 - -3 = x^2 + 2x + 4 \end{aligned}$$

This suggests that it is *possible* to have $\sqrt{-3}$ produced consistently as part of a factor of a quadratic expression. However, from our previous work we know that $\sqrt{-3}$ does not equal either $\sqrt{3}$ or $-\sqrt{3}$. So $\sqrt{-3}$ cannot be a real number, but it does exist. It would need to be part of another realm of numbers.

When we begin to consider the *existence* of $\sqrt{-3}$ we need to seek answers to questions such as:

- Are the real numbers part of a larger number system?
- Why do we limit ourselves to a one-dimensional view of numbers?
- If we expanded our view, would it be possible to see numbers as being two-dimensional?
- Is $\sqrt{-1}$ the foundation of the second dimension of numbers?

We can thank René Descartes for offering us a way to answer these questions, although he personally disregarded any answer that contained a negative number as false. However, it was his work that converted a geometric representation of mathematics into the format with which we are more familiar, an algebraic representation. He achieved this through the use of the cartesian plane, which used a grid system to locate any point on the Euclidean plane. A locus of points can therefore be described by the use of simple equations. Earlier we used the equation $y = ax^2 + bx + c$ to describe a particular conic section, the parabola.

The cartesian plane with which we are familiar uses two number lines as axes, one of which runs horizontally, the other vertically. Both use real numbers for a scale. Let us turn our attention to the vertical axis. We *could* replace its real number scale with a scale based on the square root of negative numbers. We could make the scale easier to visualise if we understood a number such as $\sqrt{-3}$ to be equivalent to $\sqrt{3 \times -1}$, which would become $\sqrt{3} \times \sqrt{-1}$. This simplification would enable us to use $\sqrt{-1}$ as the scale for the vertical axis.

This action has *removed the need* for $\sqrt{-1}$ to find a place on the real number line. In a similar way to 1 on the real number line, $\sqrt{-1}$ is the generator of all numbers on the vertical axis. The new scale represents a different set of numbers that have one thing in common: they are all multiples of $\sqrt{-1}$. That is, the vertical axis is made up of any number $a\sqrt{-1}$, where $a \in R$.

We are at a stage where we may be able to hear one mathematical hand clapping. As $\sqrt{-1}$ is *not* part of the *real* number system, we can be daring and give it a new name. It was created through our *i*magination and is beyond the boundaries laid down by the real number system, so why not just call it '*i*'. This means $\sqrt{-1} = i$ and thus $i^2 = -1$, and is analogous to the statement that if $\sqrt{2} = n$, then $n^2 = 2$. It is important to emphasise to students that finding a concrete model is an approach that provides for the *existence* of complex numbers.

We can use the imaginary axis and the real axis to form a new plane of numbers referred to as the complex plane. Jean Robert Argand was the first to publish a method of graphically representing complex numbers using the cartesian plane. Argand simply replaced the y-axis with an imaginary axis called Im z, using ai, where $a \in R$, as the scale (Figure 3.5).

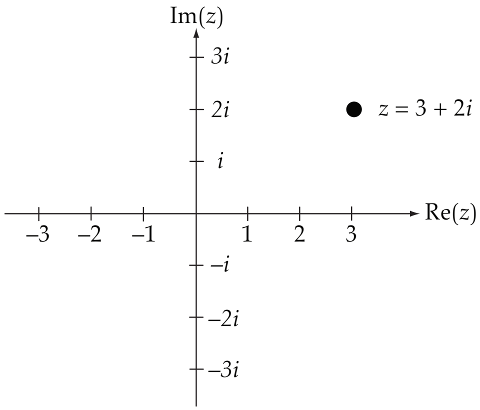


Figure 3.5: An Argand diagram plotting the complex number $z = 3 + 2i$
($z = x + yi$ where $x, y \in R$ and $i^2 = -1$)

It was through the work of mathematicians who used their imaginations, and did not allow the boundaries of their own or of the creations of others to limit their thinking, that we can come to the realisation that it is possible to visualise a 'length' that is less than zero. It may not occur in the realm of 'reality', we need to move into the world of imagination. If we concentrate and refuse to limit our view of what may be possible, we can suddenly bring into our view of the world a 'length' that can be measured as a negative.

Student Activity 3.2 will enable students to become comfortable with the manipulations involved, and thus to see that the same algebraic *processes* apply to quadratic expressions in real and complex fields.

STUDENT ACTIVITY 3.2

Find the roots of the following equations:

a $z^2 + 4z + 6 = 0$

b $z^2 + 2z + 8 = 0$

c $z^2 - 2z + 8 = 0$

d $z^2 - 4z + 10 = 0$

Student Activity 3.3 can be used to allow students to realise that a different geometric mode is used for complex numbers, while all the real numbers are included as a subset, that is, on the *real* axis.

STUDENT ACTIVITY 3.3

Plot the roots of the equations in **Student Activity 3.2** on separate Argand diagrams. Comment on what they show about any relationship between complex roots of polynomials with real coefficients.

SUMMARY

- The competition to finding the solutions of cubic polynomial equations in the sixteenth century allowed the mathematics of complex numbers to start to take form.
- Cardano was one of the first to publish a method of solving cubic equations. In his book *Ars Magna* he described Fontano's method for solving cubic equations. Students could improve their appreciation of and skills with algebra by using this method to find the solutions of cubic polynomial equations of a single variable.
- The use of complex numbers as solutions to quadratic equations is a common method used to introduce complex numbers to students in the senior secondary years. Complex solutions exist when the discriminant of a quadratic equation is less than zero.
- Complex numbers can be visualised using a special two-dimensional plane. This representation is referred to as an Argand diagram (of the complex plane). Real numbers are represented on a horizontal axis, while imaginary numbers are represented on a vertical axis. Students should be encouraged to use Argand diagrams to plot complex numbers.

References

Allendoerfer, CB & Oakley, CC 1963, *Principles of mathematics*, McGraw-Hill, New York.

Mazur, B 2003, *Imagining numbers: Particularly the square root of minus fifteen*, Penguin, London.

Websites

http://www-groups.dcs.st-and.ac.uk/~history/reference/Cardan.html
An extensive and searchable archive covering famous mathematicians and the concepts they developed.

http://www-history.mcs.st-andrews.ac.uk/References/Tartaglia.html
This sites gives a list of references of books about the life and times of *Tartaglia*.

http://www-history.mcs.st-andrews.ac.uk/Mathematicians/Cardan.html
This site gives a history of Cardano and links to a number of sites that give a good overview of the lives and times of a number of mathematicians.

CHAPTER 4

FORM AND STRUCTURE: A CAREFUL EXPOSITION ON OPERATING WITH COMPLEX NUMBERS

The conflicts and arguments between Fontano and Candano over the solutions of quadratic and cubic equations were very public affairs. So much so that another young mathematician became concerned and involved from a distance. Rafael Bombelli (1526–73) believed that the animosity evident was distracting good men from their work.

He believed that if someone took the time to write a careful exposition on algebra, many of the arguments would no longer have any foundation. Lack of understanding, too, often fans the flames of discontent. Bombelli began to write a book that he believed would allow all to understand the great work of Candano. This book, *L'Algebra*, contains an explicit and thorough account of algebra. Of particular interest to us is the first explanation of the rules for the addition, subtraction and multiplication of complex numbers.

To appreciate Bombelli's endeavour, we should travel a similar path and carefully consider the four operations with complex numbers. We need to take into consideration his concerns and advice.

Bombelli, as we may, found working with complex numbers extremely challenging. He suggested that the only way to comprehend complex numbers is to concentrate and pay careful attention, otherwise you will be prone to error. It is worth quoting him directly:

> And although to many this will appear an extravagant thing, because even I held this opinion some time ago, since it appeared to me more sophistic than true, nevertheless I searched hard and found the demonstration, which will be noted below ... But let the reader apply all his strength of mind, for [otherwise] even he will find himself deceived.

In a similar vein to Bombelli, we should begin by looking at the four operations with complex numbers. If a man of his stature found these difficult to comprehend, we would be wise to travel slowly, carefully considering each operation.

Teachers can note that the approach used proceeds by an analogy with the introduction of any new number system. We begin by introducing both notion and form (models) for the numbers, and then show how arithmetic computations work with these numbers.

ADDITION

Any complex number $z = x + yi = x + iy$, where x and $y \in R$ and $i^2 = -1$ is made up of two parts. The first part, x, is called the *real* component or $\mathrm{Re}(z)$, and y is referred to as the *imaginary* component or $\mathrm{Im}(z)$.

When we have two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, we add them by summing the *real* and *imaginary* components respectively:

$$\begin{aligned} z_1 + z_2 &= x_1 + iy_1 + x_2 + iy_2 \\ &= x_1 + x_2 + i(y_1 + y_2) \end{aligned}$$

It is worth displaying this addition on an Argand diagram (complex plane), because it offers insights into the operation of adding complex numbers using a geometric model (Figure 4.1).

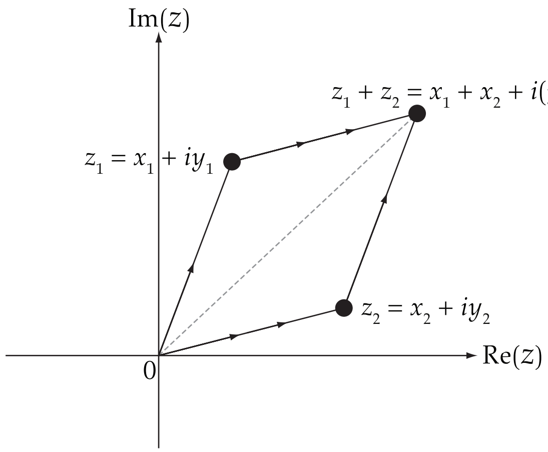


Figure 4.1: The parallelogram rule for the addition of complex numbers

This relationship is sometimes called the parallelogram rule for the addition of complex numbers, as it is based on regarding complex numbers as vectors in the complex plane. In short, the addition of two complex numbers z_1 and z_2 is $z_1 + z_2$, where $z_1 + z_2$ is the fourth vertex on a parallelogram when z_1, z_2 and the origin are the other vertices.

It would be appropriate to encourage students to become familiar with the two-dimensional nature of complex numbers by encouraging them to perform some additions that are represented on the complex plane.

STUDENT ACTIVITY 4.1

Use an Argand diagram to find the sum of the following complex numbers.

a i $2 + i$ and $3 - 2i$

ii $5 + 2i$ and $\sqrt{3} + 2i$

iii $-5 - 2i$ and $1 + i$

b i For each pair of complex numbers in part a, find the distance that each complex number and their sum are from the origin.

ii Comment on your findings

For all the student activities in this chapter it is worth noting that computer algebra systems (CAS) and many graphic calculators can perform the four arithmetic operations with complex numbers. It would be worth comparing the results gained by using technology with the methods of Bombelli.

SUBTRACTION

Subtraction of complex numbers follows a similar process to addition. When we have two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, we add them to find the difference between the real and imaginary components.

$$\begin{aligned} z_1 - z_2 &= (x_1 + iy_1) - (x_2 + iy_2) \\ &= x_1 + iy_1 - x_2 - iy_2 \\ &= x_1 - x_2 + i(y_1 - y_2) \end{aligned}$$

It might be easier to view problems of subtraction as the addition of $-1 \times z_2$. This means that $z_1 - z_2 = z_1 + -z_2$. By using this approach we will find it easier to use an Argand diagram to demonstrate the subtraction of complex numbers (Figure 4.2).

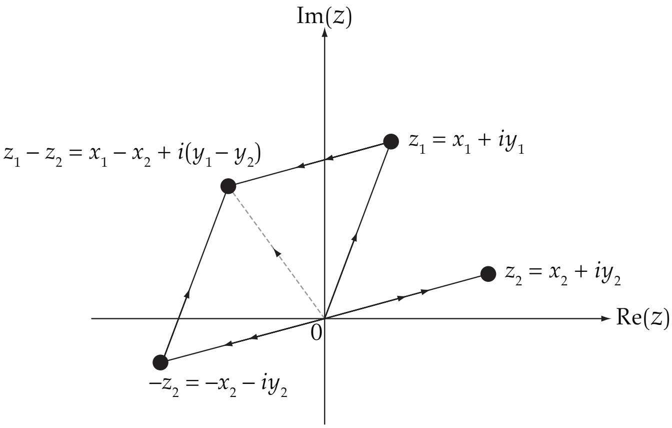


Figure 4.2: Subtraction of complex numbers represented on the complex plane

When we use an Argand diagram to view the subtraction of two complex numbers, we notice that once again a parallelogram is formed and $z_1 - z_2$ is the point on the vertex of the parallelogram which is opposite the origin.

MULTIPLICATION

We will begin by considering the multiplication of a complex number by a real number. We also use an Argand diagram to represent this multiplication (Figure 4.3).

$$az = a(x + iy) = ax + aiy \text{ when } a \in R$$

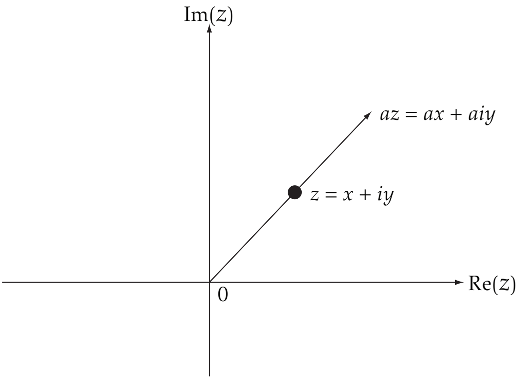


Figure 4.3: Multiplication of a complex number by a scalar represented on the complex plane

When we multiply a complex number by a real number, or scalar, the new complex number will shift either towards or away from the origin along the line that has the origin and the original complex number as two points. The real number, or scalar, multiple acts to dilate the original point on the Argand diagram.

We should now investigate the impact of multiplying a complex number by i. Let us start with the complex number $z = x + iy$.

$$iz = ix + i^2 y = -y + ix$$
$$i^2 z = -iy + i^2 x = -x - iy$$
$$i^3 z = -ix - i^2 y = y - ix$$
$$i^4 z = iy - i^2 x = x + iy = z$$

To allow a deeper understanding of what occurs when we multiply a complex number by i, we will use an Argand diagram (Figure 4.4).

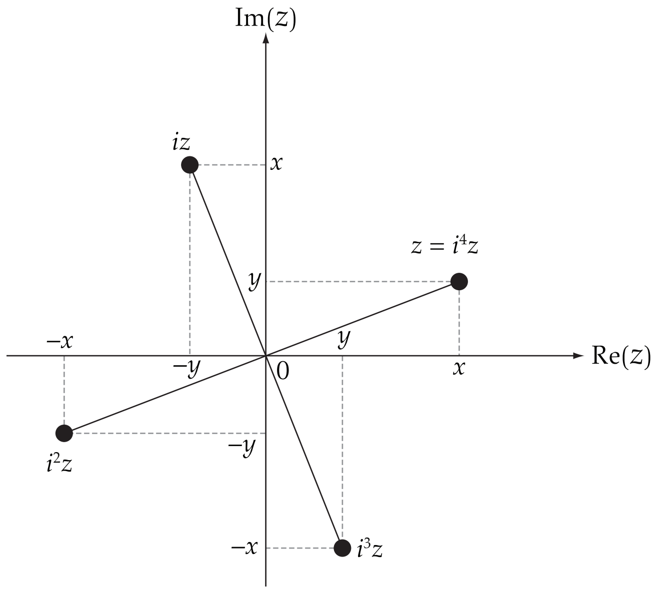


Figure 4.4: The effect of multiplying a complex number by *i* represented on the complex plane

Two things become evident on the Argand diagram. The first is that by Pythagoras' theorem it is clear that each point is the same distance from the origin. The second observation is that when we multiply a complex number by i the new complex number is rotated about the origin in an anticlockwise direction. Upon closer inspection, it is clear that each rotation is equivalent, and after the fourth rotation we return to our starting point. This would suggest that multiplying a complex number by i will give a new complex number which is the same distance from the origin but has been rotated 90° anticlockwise. Thus teachers will find that the language of transformations of the plane is useful here.

Multiplication of complex numbers of the form $z = x + iy$ is defined as follows, using the distribution property, and $i^2 = -1$:

Let $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$

$$\begin{aligned} z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\ &= x_1 x_2 + x_1 iy_2 + iy_1 x_2 + i^2 y_1 y_2 \\ &= x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1) \end{aligned}$$

Students should be encouraged to explore in some depth the properties of the multiplication of complex numbers. An activity such as Student Activity 4.2 allows students to investigate the transformation properties of multiplication of complex numbers.

STUDENT ACTIVITY 4.2

Consider the complex numbers $z_1 = 1 + i$, $z_2 = 1 + \sqrt{3}i$ and $z_3 = \sqrt{3} + i$.

a Evaluate $z_1 z_2$, $z_1 z_3$ and $z_2 z_3$.

b i Plot z_1, z_2 and $z_1 z_2$ on an Argand diagram.
  ii Find the distance from the origin to each of z_1, z_2 and $z_1 z_2$.
  iii Find the angles between the Re(z)-axis and the line intervals that join the points z_1, z_2 and $z_1 z_2$ to the origin.

c i Plot z_1, z_3 and $z_1 z_3$ on an Argand diagram.
  ii Find the distance from the origin to each of z_1, z_3 and $z_1 z_3$.
  iii Find the angles between the Re(z)-axis and the line intervals that join the points z_1, z_3 and $z_1 z_3$ to the origin.

d i Plot z_2, z_3 and $z_2 z_3$ on an Argand diagram.
  ii Find the distance from the origin to each of z_2, z_3 and $z_2 z_3$.
  iii Find the angles between the Re(z)-axis and the line intervals that join the points z_2, z_3 and $z_2 z_3$ to the origin.

e Comment on what your results indicate about the multiplication of complex numbers.

POLAR COORDINATES

The cartesian coordinate system is invaluable for showing us where a point is located on the plane. However, it is not very informative about what path we should use to reach that point, unless we want to travel a certain distance along the x-axis and then turn through 90º and travel another given distance. To arrive at a point in the most efficient way, we need a system that points the direction and then tells us how far we should travel in that direction. This is the notion that lies behind a vector for airline navigation. The vector tells the pilot to travel on a particular compass bearing. After the airliner has travelled a set distance on this path it should arrive at its destination.

We can use a similar idea to locate points on the complex plane. We could inform others of its location on the plane by giving the distance of the point from the origin and the direction of travel from the origin to the point. We refer to the distance as the magnitude, r, or modulus and the rotation about the origin as the phase, θ. So a point (x, y) would hold the same position as the point $[r, \theta]$. When we use $[r, \theta]$ to locate a point on the plane, we are using polar coordinates.

It is not difficult to convert cartesian coordinates to polar coordinates and vice versa. From Figure 4.5, we notice that we can use Pythagoras' theorem to establish the value of r, where $r^2 = x^2 + y^2 \Rightarrow r = \sqrt{x^2 + y^2}$.

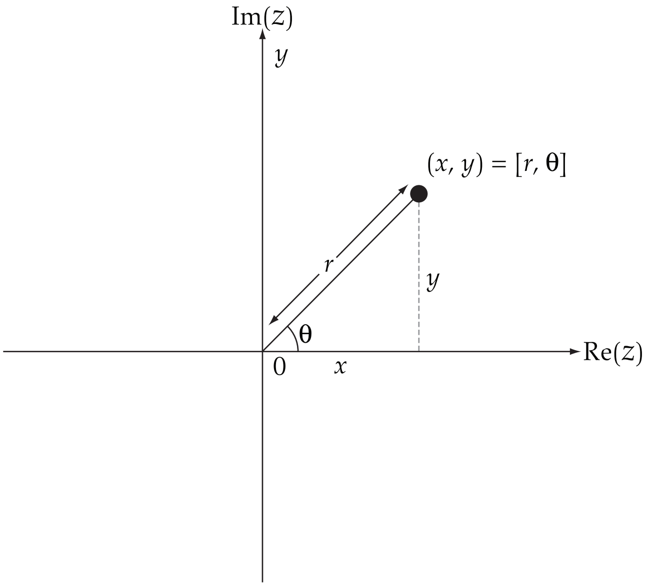


Figure 4.5: The relationship between cartesian and polar coordinates

The value of θ can be established by trigonometry and circular functions.

$$\theta = \tan^{-1}\left(\frac{y}{x}\right)$$

The value of θ is called the *argument* of the complex number z and written $\arg(z)$. We can further define a unique value $-\pi < \arg(z) \leq \pi$ to avoid multiple values due to the periodic nature of the circular functions and hence $\arg(z)$.

If we have been given the polar coordinates of a point, we can use trigonometry to find the equivalent cartesian coordinates.

$$x = r\cos(\theta) \text{ and } y = r\sin(\theta)$$

We now have two methods of locating a point on an Argand diagram.

$$z = x + iy \text{ and } z = r\cos(\theta) + ri\sin(\theta) = r(\cos(\theta) + i\sin(\theta))$$

For the sake of brevity, $z = r\cos(\theta) + ri\sin(\theta) = r(\cos(\theta) + i\sin(\theta))$ is usually referred to as $z = r\text{cis}(\theta)$.

An interesting aspect of the use of polar coordinates for the location of points on the complex plane is that they use rotation from an axis and distance from an origin. This could prove to be valuable when we consider the multiplication of complex numbers because, when complex numbers are multiplied, points are both dilated from and rotated about the origin. This should encourage us to multiply two complex numbers expressed in polar form.

Let $z_1 = r_1(\cos(\theta_1) + i\sin(\theta_1))$ and $z_2 = r_1(\cos(\theta_2) + i\sin(\theta_2))$.

$$\begin{aligned}
z_1 z_2 &= r_1(\cos(\theta_1) + i\sin(\theta_1))r_2(\cos(\theta_2) + i\sin(\theta_2)) \\
&= r_1 r_2(\cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2) + i\sin(\theta_1)\cos(\theta_2) + i\sin(\theta_2)\cos(\theta_1)) \\
&= r_1 r_2((\cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2)) + i(\sin(\theta_1)\cos(\theta_2) + \sin(\theta_2)\cos(\theta_1)))
\end{aligned}$$

Using the compound angle formula for the circular functions:

$$\cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2) = \cos(\theta_1 + \theta_2) \text{ and}$$
$$\sin(\theta_1)\cos(\theta_2) + \sin(\theta_2)\cos(\theta_1) = \sin(\theta_1 + \theta_2)$$

gives:

$$\begin{aligned}
z_1 z_2 &= r_1 r_2((\cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2)) + i(\sin(\theta_1)\cos(\theta_2) + \sin(\theta_2)\cos(\theta_1))) \\
&= r_1 r_2(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)) \\
&= r_1 r_2\,\text{cis}(\theta_1 + \theta_2)
\end{aligned}$$

This result can also be interpreted in terms of compositions of transformations—two dilations and two rotations. We now have two approaches to the multiplication of complex numbers:

- using polar coordinates $z_1 z_2 = r_1 r_2 \operatorname{cis}(\theta_1 + \theta_2)$
- using cartesian coordinates $z_1 z_2 = x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1)$

At this point it would be worth revisiting Student Activity 4.2, this time using the polar form of a complex number, so that students can appreciate that while there are two approaches to the multiplication of complex numbers they give the same results.

STUDENT ACTIVITY 4.3

Consider the complex numbers $z_1 = 1 + i$, $z_2 = 1 + \sqrt{3}i$ and $z_3 = \sqrt{3} + i$.

a Express z_1, z_2 and z_3 in polar form.

b Evaluate $z_1 z_2$, $z_1 z_3$ and $z_2 z_3$ using the polar form.

c Plot z_1, z_2 and $z_1 z_2$ on an Argand diagram.

d Plot z_1, z_3 and $z_1 z_3$ on an Argand diagram.

e Plot z_2, z_3 and $z_2 z_3$ on an Argand diagram.

DIVISION

When we divide one complex number by another we are attempting to simplify the expression

$$\frac{z_1}{z_2} = \frac{x_1 + iy_1}{x_2 + iy_2}$$

For example if $z_1 = 1 + i$ and $z_2 = 2 + 3i$ then $\frac{z_1}{z_2} = \frac{1 + i}{2 + 3i}$. The question with this division is where to start. The first step is not as intuitive as the first step in the previous three operations. For a hint we need to revisit the problem that led us to imagine a solution to the square root of a negative number. We were asked to fully factorise the quadratic equation $x^2 + 2x + 4$. The solution we obtained was $(x + 1 - \sqrt{-3})(x + 1 + \sqrt{-3})$, which we now know is $(x + 1 - i\sqrt{3})(x + 1 + i\sqrt{3})$.

This means that $x^2 + 2x + 4 = (x + 1 - i\sqrt{3})(x + 1 + i\sqrt{3})$.

What happens to this equation when the value of x is equal to zero? The right-hand side of the equation would be $(1 - i\sqrt{3})(1 + i\sqrt{3})$. This is the difference of two squares $(a - b)(a + b) = a^2 - b^2$. In this case $a = 1$, and $b = i\sqrt{3}$.

$$\begin{aligned}(a - b)(a + b) &= a^2 - b^2 \\ (1 - i\sqrt{3})(1 + i\sqrt{3}) &= 1^2 - (i\sqrt{3})^2 \\ &= 1 - i^2 \times 3 \\ &= 1 + 3 \\ &= 4\end{aligned}$$

This is as we would expect, because when $x = 0$, $x^2 + 2x + 4 = 4$.

What is important to us is to recognise that the product of these complex numbers is a real number. We can generalise this result for all complex numbers. When $z_1 = x + iy$ is multiplied by the complex number $z_2 = x - iy$, the result will be a positive real number.

$$\begin{aligned}z_1 z_2 &= (x + iy)(x - iy) \\ &= x^2 - i^2 y^2 \\ &= x^2 + y^2\end{aligned}$$

This relationship between z_1 and z_2 is an important property of complex numbers: z_1 and z_2 are called a *conjugate pair*. We can create a *conjugate* of any complex number by reversing the sign of the imaginary component of the number. The conjugate is referred to as $\bar{z}$ or 'zed bar'.

If $z = x + iy$, then $\bar{z} = x - iy$.

If $z = x - iy$, then $\bar{z} = x + iy$.

Student Activity 4.4 provides the opportunity for students to explore some of the properties of complex numbers and their conjugates in cartesian form.

STUDENT ACTIVITY 4.4

Prove the following properties of the complex conjugate.

1. $\bar{\bar{z}} = z$
2. $\bar{z} - z = 2yi$
3. $z\bar{z} = x^2 + y^2$
4. $z + \bar{z} = 2x$
5. $\overline{u + v} = \bar{u} + \bar{v}$
6. $\overline{u - v} = \bar{u} - \bar{v}$
7. $\overline{uv} = \bar{u}\,\bar{v}$
8. $z \in R$ if and only if $z = \bar{z}$

We can now return to the division $\frac{1+i}{2+3i}$. We can simplify this problem by multiplying the numerator and denominator by the conjugate of the denominator.

$$\begin{aligned}\frac{1+i}{2+3i}\times\frac{2-3i}{2-3i} &= \frac{2+3+2i-3i}{4+9}\\ &= \frac{5-i}{13}\\ &= \frac{5}{13}-\frac{i}{13}\end{aligned}$$

In general terms:

$$\begin{aligned}\frac{z_1}{z_2} &= \frac{x_1+iy_1}{x_2+iy_2}\\ &= \frac{x_1+iy_1}{x_2+iy_2}\times\frac{x_2-iy_2}{x_2-iy_2}\\ &= \frac{x_1x_2+y_1y_2+i(x_2y_1-x_1y_2)}{x_2^2+y_2^2}\end{aligned}$$

This process is analogous to the rationalisation of a real surd expression. Student Activity 4.5 allows students to explore not only the operation of complex number divisions, but also some of its geometric properties.

STUDENT ACTIVITY 4.5

Consider the complex numbers $z_1 = 1 + i$, $z_2 = 1 + \sqrt{3}i$ and $z_3 = \sqrt{3} + i$.

a Evaluate $\frac{z_1}{z_2}$, $\frac{z_1}{z_3}$ and $\frac{z_2}{z_3}$.

b i Plot z_1, z_2 and $\frac{z_1}{z_2}$ on an Argand diagram.

ii Find the distance from the origin to each of z_1, z_2 and $\frac{z_1}{z_2}$.

iii Find the angles between the Re(z)-axis and the line intervals that join the points z_1, z_2 and $\frac{z_1}{z_2}$ to the origin.

c i Plot z_1, z_3 and $\frac{z_1}{z_3}$ on an Argand diagram.

ii Find the distance from the origin to each of z_1, z_3 and $\frac{z_1}{z_3}$.

iii Find the angles between the Re(z)-axis and the line intervals that join the points z_1, z_3 and $\frac{z_1}{z_3}$ to the origin.

d i Plot z_2, z_3 and $\frac{z_2}{z_3}$ on an Argand diagram.

ii Find the distance from the origin to each of z_2, z_3 and $\frac{z_2}{z_3}$.

iii Find the angles between the Re(z)-axis and the line intervals that join the points z_2, z_3 and $\frac{z_2}{z_3}$ to the origin.

We can also use the polar forms of complex numbers to divide two complex numbers. To do this we need to be able to find the conjugate of $z = rcis(\theta)$. We can do this if we have an understanding of the symmetry of the circular functions $y = \sin(x)$ and $y = \cos(x)$. This can best be achieved by studying the graphs of these functions (Figures 4.6 and 4.7).

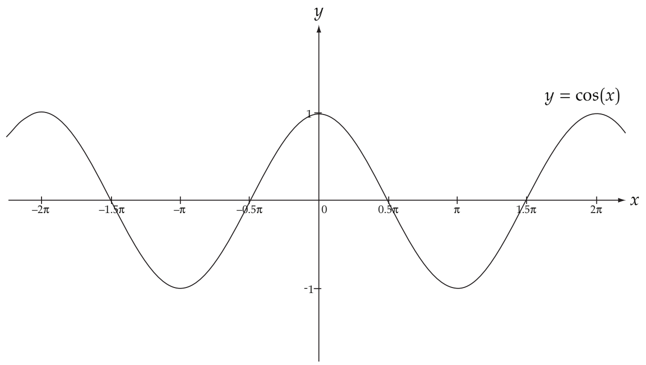


Figure 4.6: The graph of $y = \cos(x)$

We notice that the graph of $y = \cos(x)$ is symmetrical about the y-axis, which means that for any value of x, $\cos(-x) = \cos(x)$.

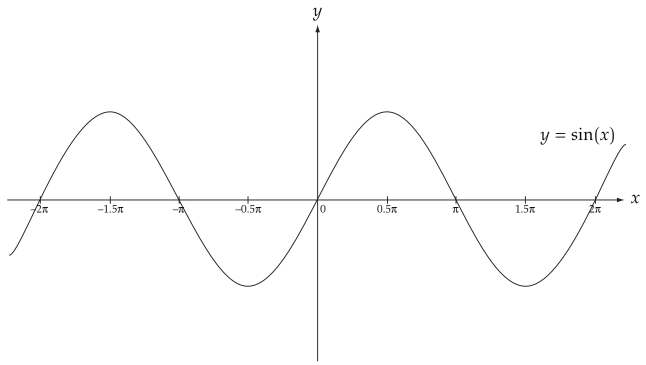


Figure 4.7: The graph of $y = \sin(x)$

Notice that the graph of $y = \sin(x)$ has half-turn rotational symmetry about the origin, which means that for any value of x, $\sin(-x) = -\sin(x)$.

We can use these properties of circular functions to find the conjugate of a complex number when it is expressed in polar form.

The expansion of $z = r\text{cis}(\theta)$ is $z = r(\cos(\theta) + i\sin(\theta))$. The conjugate of $z = r(\cos(\theta) + i\sin(\theta))$ would be $\bar{z} = r(\cos(\theta) - i\sin(\theta))$. Using the properties $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$, we can change $\bar{z} = r(\cos(\theta) - i\sin(\theta))$ to become $\bar{z} = r(\cos(-\theta) + i\sin(-\theta))$, which is equivalent to $\bar{z} = r\text{cis}(-\theta)$.

So the conjugate pairings for complex numbers expressed in polar form are:

- $z = r\text{cis}(\theta)$ and $\bar{z} = r\text{cis}(-\theta)$
- $z = r\text{cis}(-\theta)$ and $\bar{z} = r\text{cis}(\theta)$

We can use the polar form to find the quotient of two complex numbers:

$$\begin{aligned}\frac{z_1}{z_2} &= \frac{r_1\text{cis}(\theta_1)}{r_2\text{cis}(\theta_2)}\\ &= \frac{r_1\text{cis}(\theta_1)}{r_2\text{cis}(\theta_2)} \times \frac{r_2\text{cis}(-\theta_2)}{r_2\text{cis}(-\theta_2)}\\ &= \frac{r_1}{r_2} \times \frac{\text{cis}(\theta_1 - \theta_2)}{\cos^2(\theta_2) + \sin^2(\theta_2)}\\ &= \frac{r_1}{r_2} \times \frac{\text{cis}(\theta_1 - \theta_2)}{1}\\ &= \frac{r_1}{r_2}\text{cis}(\theta_1 - \theta_2)\end{aligned}$$

This result can also be interpreted in terms of compositions of transformations—a dilation and an inverse dilation, a rotation and an inverse rotation.

Student Activity 4.6 allows students to repeat the divisions they performed in Student Activity 4.5, this time using the polar form of complex numbers. This should allow them to realise that both forms of division produce the same result.

STUDENT ACTIVITY 4.6

Consider the complex numbers $z_1 = 1 + i$, $z_2 1 + \sqrt{3}i$ and $z_3 = \sqrt{3} + i$.

a Express z_1, z_2 and z_3 in polar form.

b Evaluate $\frac{z_1}{z_2}$, $\frac{z_1}{z_3}$ and $\frac{z_2}{z_3}$ using polar form.

c Plot z_1, z_2 and $\frac{z_1}{z_2}$ on an Argand diagram.

d Plot z_1, z_3 and $\frac{z_1}{z_3}$ on an Argand diagram.

e Plot z_2, z_3 and $\frac{z_2}{z_3}$ on an Argand diagram.

SUMMARY

- Bombelli gave a sound description of the four arithmetic operations on complex numbers.
  - **Addition**

$$\begin{aligned} z_1 + z_2 &= x_1 + iy_1 + x_2 + iy_2 \\ &= x_1 + x_2 + i(y_1 + y_2) \end{aligned}$$

  - **Subtraction**

$$\begin{aligned} z_1 - z_2 &= (x_1 + iy_1) - (x_2 + iy_2) \\ &= x_1 + iy_1 - x_2 - iy_2 \\ &= x_1 - x_2 + i(y_1 - y_2) \end{aligned}$$

  - **Multiplication**

$$kz = k(x + iy) = kx + kyi \text{ for } k \in R$$

$$\begin{aligned} z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\ &= x_1 x_2 + x_1 iy_2 + iy_1 x_2 + i^2 y_1 y_2 \\ &= x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1) \end{aligned}$$

  - **Division**
    Division of complex numbers relies on the use of the complex conjugate.

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{x_1 + iy_1}{x_2 + iy_2} \\ &= \frac{x_1 + iy_1}{x_2 + iy_2} \times \frac{x_2 - iy_2}{x_2 - iy_2} \\ &= \frac{x_1 x_2 + y_1 y_2 + i(x_2 y_1 - x_1 y_2)}{(x_2)^2 + (y_2)^2} \end{aligned}$$

- Polar coordinates can also be used to represent complex numbers. This representation of complex numbers makes the multiplication of complex number easier. Using Polar coordinates

$$z_1 z_2 = r_1 r_2 \operatorname{cis}(\theta_1 + \theta_2) \text{ and } \frac{z_1}{z_2} = \frac{r_1}{r_2} \operatorname{cis}(\theta_1 - \theta_2).$$

Further reading

Crossely, JN 1987, *The emergence of number*, World Scientific, Singapore.

Websites

http://www-history.mcs.st-andrews.ac.uk/Mathematicians/Bombelli.html
This website contains a brief history of Bombelli, with links to other sites for further biographies and explanations.

http://nsm1.nsm.iup.edu/gsstoudt/history/bombelli/bombelli.pdf
This website contains a version of Bombelli's original work with some explanations that help the reader with a more current understanding of mathematics.

CHAPTER 5

THE GENIUS OF GAUSS

Carl Friedrich Gauss was considered by many to be the greatest mathematician of the 19th century. He once said 'Mathematics is the queen of the sciences and number theory is the queen of mathematics'. A little less all-encompassing than the Pythagorean School's 'all is number', but it does tell us something of Gauss's belief in the power of number and thus of mathematics.

Gauss belonged to the 19th century, the century in which mathematics gave us non-Euclidean geometry and n-dimensional spaces. For Western mathematics it was possibly the second most revolutionary and productive century in its long history. It was almost as productive as the 5th century BC in which the mathematicians of Greek civilisation changed the way we 'imagined' the mathematical world by proving, among so many other things, that we needed to add irrational numbers to the field of real numbers.

It is enough to explore two aspects of Gauss's work to appreciate his genius: the Fundamental Theorem of Algebra and the Fundamental Theorem of Arithmetic. Teachers could introduce both of these theorems to students as an extension of previous work with factors and factorisation to indicate their importance in terms of the generalisation of these properties and for conceptual completeness.

THE FUNDAMENTAL THEOREM OF ALGEBRA

Gauss is believed to be the first to prove the Fundamental Theorem of Algebra. He did this in his doctoral thesis of 1799 which was entitled 'New demonstrations of the theorem that every rational integral algebraic function in one variable can be resolved into real factors of first and second degree'. We are perhaps more familiar with another wording of the Fundamental Theorem of Algebra:

> Every polynomial with real coefficients can be expressed as the product of real linear and real quadratic factors.

This can be taken a step further by entertaining the possibility that a polynomial may have complex coefficients:

> Every polynomial of degree *n* with complex coefficients has *n* roots that can be expressed as complex numbers.

Expressing this mathematically:

$$a_n z^n + a_{n-1} z^{n-1} + a_{n-2} z^{n-2} + a_{n-3} z^{n-3} + a_{n-4} z^{n-4} \ldots + a_1 z^1 + a_0$$
$$= a(z - b_n)(z - b_{n-1})(z - b_{n-2})\ldots(z - b_1)$$

where a_j and b_j, $j \in N$, are complex numbers. This means that $z = b_1, b_2, b_3 \ldots b_n$ are all roots of the original expression.

We can use the Fundamental Theorem of Algebra to explore some interesting ideas that will allow us to deepen our understanding of factorisation and the complex number field.

We should begin this exploration by considering as an example, the polynomial $P(z) = z^3 + 4z^2 + 7z + 4$, which has real coefficients. Using the Fundamental Theorem of Algebra we know that $P(z) = (z + a)(z + b)(z + c)$, which means that $z^3 + 4z^2 + 7z + 4 \equiv (z + a)(z + b)(z + c)$. We can use the notation 'identically equal to ($\equiv$)' because both sides of the equation will remain equal regardless of any value chosen for z. This is an equivalence relation.

If we let $z = 0$, then:

$$0^3 + 4 \times 0^2 + 7 \times 0 + 4 = (0 + a)(0 + b)(0 + c)$$
$$4 = abc$$

This suggests that a, b and c are all factors of 4.

We also know that if $P(z) = 0$ then $(z + a)(z + b)(z + c) = 0$.

From the Null Factor Law, we know that $z = -a, -b$ or $-c$, so $P(-a) = P(-b) = P(-c) = 0$. As a consequence, if we can find a value d such that $P(d) = 0$, then $z - d$ must be a factor of $P(z)$.

Putting all the above together, we have a method of searching for factors of a polynomial. For $P(z) = z^3 + 4z^2 + 7z + 4$, we could use the most obvious factors of 4 to find the first root of $P(z)$. The integer factors of 4 are $\pm 1, \pm 2$, and ± 4.

$P(1) = 1^3 + 4 \times 1^2 + 7 \times 1 + 4 \neq 0$, therefore 1 is not a root of $P(z)$.

$P(-1) = (-1)^3 + 4 \times (-1)^2 + 7 \times (-1) + 4 = 0$, therefore –1 is a root of $P(z)$, and $z + 1$ is a factor.

To find the remaining factor we will use division of polynomials.

$$\begin{array}{r} z^2 + 3z + 4 \\ z+1\overline{\big)\, z^3 + 4z^2 + 7z + 4} \\ z^3 + z^2 \qquad\qquad\quad \\ 3z^2 + 7z + 4 \\ 3z^2 + 3z \qquad \\ 4z + 4 \\ 4z + 4 \\ 0 + 0 \end{array}$$

This means that:

$$\frac{z^3 + 4z^2 + 7z + 4}{z + 1} = z^2 + 3z + 4$$
$$\Rightarrow z^3 + 4z^2 + 7z + 4 = (z + 1)\left(z^2 + 3z + 4\right)$$

Alternatively, the method of equating coefficients can be used to obtain:

$$\begin{aligned} P(z) &= z^3 + 4z^2 + 7z + 4 \\ &= z^2(z + 1) + 3z(z + 1) + 4(z + 1) \\ &= (z + 1)\left(z^2 + 3z + 4\right) \end{aligned}$$

In this case we notice that $P(z) = z^3 + 4z^2 + 7z + 4$ has been expressed as a real linear factor, $z + 1$, and a real quadratic factor, $z^2 + 3z + 4$, as indicated by the Fundamental Theorem of Algebra.

We should now turn our attention to the quadratic factor $z^2 + 3z + 4$. This can be factorised by using the quadratic formula, or by completing the square.

$$\begin{aligned} z^2 + 3z + 4 &= \left(z + \frac{3 + \sqrt{-7}}{2}\right)\left(z + \frac{3 - \sqrt{-7}}{2}\right) \\ &= \left(z + \frac{3 + i\sqrt{7}}{2}\right)\left(z + \frac{3 - i\sqrt{7}}{2}\right) \end{aligned}$$

So $z^3 + 4z^2 + 7z + 4 = (z + 1)\left(z + \frac{3 + i\sqrt{7}}{2}\right)\left(z + \frac{3 - i\sqrt{7}}{2}\right)$

When expressed as the roots of the equations, the results would be:

$$z = 1, -\frac{3 + i\sqrt{7}}{2} \text{ and } -\frac{3 - i\sqrt{7}}{2}$$

The original polynomial was of degree 3, and thus, according to the Fundamental Theorem of Algebra, we should expect three complex roots.

While it may appear that we are stating the obvious, we need to realise that 1 *is* a complex number; it can be written as $1 + 0i$.

We can obtain results using technology. Exact answers can be obtained by using equation-solving functions of computer algebra systems, and programs that will provide roots for polynomial equations are freely available for graphics calculators. This family of graphics calculator programs will give roots and factors and perform division of polynomials.

It is worth noting that $-\frac{3 + i\sqrt{7}}{2}$ and $-\frac{3 - i\sqrt{7}}{2}$ are a conjugate pair.

We would expect a conjugate pair when factorising a quadratic, given the nature of the quadratic formula.

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

When $b^2 - 4ac < 0$, the discriminant Δ is negative. Rewriting the quadratic formula for $\Delta < 0$:

$$\begin{aligned} z &= \frac{-b \pm \sqrt{-\Delta}}{2a} \\ &= \frac{-b \pm i\sqrt{\Delta}}{2a} \\ &= \frac{-b + i\sqrt{\Delta}}{2a} \text{ and } \frac{-b - i\sqrt{\Delta}}{2a} \end{aligned}$$

We can return to the Fundamental Theorem of Algebra with this result in mind. It states that every polynomial with ***real*** coefficients can be expressed as the product of ***real*** linear and ***real*** quadratic factors.

When the real quadratic factors of a polynomial are reduced to linear factors, these factors will be conjugate pairs. This property of the factorisation of polynomials with real coefficients is called the Conjugate Root Theorem.

CONJUGATE ROOT THEOREM

The non-real roots of a polynomial expression that has only real coefficients occur as conjugate pairs. Some students believe that the conjugate root theorem applies to all complex polynomials. Teachers could combine Student Activity 5.1 with examples in which the conjugate root theorem does not apply.

STUDENT ACTIVITY 5.1

1 Show that $\overline{z^2} = \overline{z}^2$ and $\overline{z^3} = \overline{z}^3$ for $3 + 4i$ and $2 - 3i$.

2 Show that $\overline{z^2} = \overline{z}^2$ and $\overline{z^3} = \overline{z}^3$ for $z = x + yi$. Hint use $\overline{uv} = \overline{u}\,\overline{v}$

3 Consider a general polynomial of degree n with real coefficients $P_n(z)$, where $n > 1$.

a Show that if w is a factor of $P_n(z)$ then $\overline{w}$ is also a factor.

b Show that $(z - w)(z - \overline{w})$ is a quadratic with real coefficients.

c Hence, show that if w is not real and is a root of $P_n(z)$ then $P_n(z) = (z - w)(z - \overline{w})P_{n-2}(z)$.

d Use your result in part c to prove the conjugate root theorem.

An interesting aspect of the Fundamental Theorem of Algebra is the factorisation of polynomials of the form $z^n - a^n$. The theorem suggests that we should expect n roots.

Consider, for example, $z^6 - 64$.

Using the difference of two squares:

$$(z^3 - 8)(z^3 + 8)$$

Using the difference of two cubes:

$$(z - 2)(z^2 + 2z + 4)(z^3 + 8)$$

Using the sum of two cubes:

$$(z - 2)(z^2 + 2z + 4)(z + 2)(z^2 - 2z + 4)$$

Using the quadratic formula:

$$(z - 2)(z + 1 + i\sqrt{3})(z + 1 - i\sqrt{3})(z + 2)(z - 1 + i\sqrt{3})(z - 1 - i\sqrt{3})$$

So the roots of $z^6 - 64$ are $z = \pm 2, 1 \pm i\sqrt{3}$. It is worth noting that the non-real roots occur in conjugate pairs, as is suggested by the conjugate root theorem.

Student Activity 5.2 enables students to realise that these roots have a geometric representation.

STUDENT ACTIVITY 5.2

a Plot each of the roots of $z^6 - 64$ onto an Argand diagram.

b Find the distance between each point and the origin.

c Find the angle between the x-axis and the line interval that joins each point to the origin.

d Comment on your findings.

It is possible to solve equations of the form $z^n - a^n$ where $n \in N$ *and* $a \in R$ using the polar form of complex numbers. This can be demonstrated by starting with the complex number $z = r\text{cis}(\theta)$.

We will start by raising z to a series of powers beginning with z^2.

$$z^2 = r \times r\text{cis}(\theta + \theta) = r^2\text{cis}(2\theta)$$
$$z^3 = r \times r^2\text{cis}(2\theta + \theta) = r^3\text{cis}(3\theta)$$
$$z^4 = r \times r^3\text{cis}(3\theta + \theta) = r^4\text{cis}(4\theta)$$

This pattern continues to $z^n = r^n\text{cis}(n\theta)$.

This would also suggest that $\sqrt[n]{z} = z^{\frac{1}{n}} = r^{\frac{1}{n}}\text{cis}\left(\frac{\theta}{n}\right) = \sqrt[n]{r}\,\text{cis}\left(\frac{\theta}{n}\right)$. However, this would only give one result, so we need to extend this equation to account for all the solutions:

$$z^{\frac{1}{n}} = r^{\frac{1}{n}}\text{cis}\left(\frac{2k\pi + \theta}{n}\right) \text{ where } k \in N$$

We refer to $z^n = r^n\text{cis}(n\theta)$ and $z^{\frac{1}{n}} = r^{\frac{1}{n}}\text{cis}\left(\frac{2k\pi + \theta}{n}\right)$ as De Moivre's theorem.

De Moivre's theorem can be used to solve $z^6 - 64 = 0$.

$$\begin{aligned} z^6 - 64 &= 0 \\ z^6 &= 64 \\ &= 64\text{cis}(0) \\ z &= \sqrt[6]{64}\,\text{cis}\left(\frac{2k\pi + 0}{6}\right) \end{aligned}$$

For $k = 0$: $z = 2\text{cis}(0)$
$= 2$

For $k = 1$: $z = 2\text{cis}\left(\frac{2\pi}{6}\right)$
$= 2\text{cis}\left(\frac{\pi}{3}\right)$

For $k = 2$: $z = 2\text{cis}\left(\frac{6\pi}{6}\right)$
$= 2\text{cis}(\pi)$
$= -2$

For $k = 4$: $z = 2\text{cis}\left(\frac{8\pi}{6}\right)$

$= 2\text{cis}\left(\frac{4\pi}{3}\right)$

We need to remember that $-\pi < \text{Arg}(z) \leq \pi$, which means, in this case, that $z = 2\text{cis}\left(\frac{-2\pi}{3}\right)$.

For $k = 5$: $z = 2\text{cis}\left(\frac{10\pi}{6}\right)$

$= 2\text{cis}\left(\frac{5\pi}{3}\right)$

$= 2\text{cis}\left(\frac{-\pi}{3}\right)$

For $k = 6$: $z = 2\text{cis}\left(\frac{12\pi}{6}\right)$

$= 2\text{cis}(2\pi)$

$= 2$

Notice that when $k = 6$, we obtain the same result as when $k = 0$. We have returned to our starting point and have six results, as is suggested by the Fundamental Theorem of Algebra.

As we would expect, the cartesian and polar forms of a complex number give the same results when we solve the equation $z^6 - 64 = 0$. In each case we have six roots, as is suggested by the Fundamental Theorem of Algebra. With his theorem De Moivre has given us an easy way of solving any equation of the form $z^n - a^n = 0$. Gauss used the work of De Moivre to prove the Fundamental Theorem of Algebra for a second time. In fact, Gauss offered four separate proofs of this theorem; the final proof was offered on the 50th anniversary of his first proof. Some things are worth celebrating.

Given the general nature of the Fundamental Theorem of Algebra and its proof, it would be appropriate to encourage students to explain some of the generalisations that have been hinted at in the previous work. Some of the explanations are required of students in Student Activity 5.3.

STUDENT ACTIVITY 5.3

1 Why are there n solutions for $z^n = r\text{cis}(\theta)$?
2 Why, when plotted on an Argand diagram, are the solutions to $z^n = r\text{cis}(\theta)$ equally spaced around a circle?
3 Show that if $z = \text{cis}(\theta)$ is solution to $z^n = b$, where b is a real number, then $\overline{z}$ is also a solution.

THE FUNDAMENTAL THEOREM OF ARITHMETIC

The Fundamental Theorem of Arithmetic states that every integer greater than 1 can be expressed as the product of prime numbers in one way and one way only. A prime number is a number which has two distinct factors, itself and 1. This means that 2, 3, 5, 7, 11, 13 and 17 are prime numbers, but 1, 4, 9, 12 and 15 are not.

Using the Fundamental Theorem of Arithmetic, we can express the following as the product of prime numbers (or 'primes').

$$\begin{aligned} 60 &= 2 \times 30 \\ &= 2 \times 2 \times 15 \\ &= 2 \times 2 \times 3 \times 5 \\ &= 2^2 \times 3 \times 5 \end{aligned}$$

$$\begin{aligned} 1000 &= 2 \times 500 \\ &= 2 \times 2 \times 250 \\ &= 2 \times 2 \times 2 \times 125 \\ &= 2 \times 2 \times 2 \times 5 \times 25 \\ &= 2 \times 2 \times 2 \times 5 \times 5 \times 5 \\ &= 2^3 \times 5^3 \end{aligned}$$

In both cases there are no other factorisations of 60 and 1000 that involve the products of prime numbers.

The Fundamental Theorem of Arithmetic informs us of the importance of prime numbers to number theory. They are the building blocks of all the positive integers. Gauss wondered if primes also existed in the complex field.

It is a relatively simple matter to extend integers into the complex field by first realising that any integer is already a member of the field. All integers can be expressed as $m + ni$, where $m \in Z$ and $n = 0$. For the sake of clarity we will call this group of integers the ordinary integers. Gauss extended the set of integers by allowing both m and ni to be integers. Any integer of this form is called a Gaussian integer.

The first question we can pose with the introduction of Gaussian integers is: do all the primes of the ordinary integers remain prime in the complex field, C? In fact some of the ordinary primes are no longer prime when we use Gaussian integers. For example $5 = (1 - 2i)(1 + 2i) = 1^2 + 2^2$.

This result suggests that we will need to distinguish between primes of positive integers and those of Gaussian integers. We will call this latter group the Gaussian primes. Thus a Gaussian prime is defined as a Gaussian integer with modulus greater than 1 which is not the product of Gaussian integers of smaller modulus.

We defined the modulus $|m + ni|$ of a complex number $m + ni$ to be $\sqrt{m^2 + n^2}$. This result was used when we were converting complex numbers in cartesian form into their polar equivalent. The modulus is the distance of a complex number from the origin when it is plotted on an Agrand diagram.

If we return to our work with conjugate pairs, we realise that:

$$(m - ni)(m + ni) = m^2 + n^2$$

This suggests that any ordinary prime that is the sum of two square integers will not be a Gaussian prime.

Some examples are:

$$\begin{aligned} 2 &= 1^2 + 1^2 = (1 + i)(1 - i) \\ 13 &= 2^2 + 3^2 = (2 + 3i)(2 - 3i) \\ 17 &= 4^2 + 1^2 = (4 + i)(4 - i) \end{aligned}$$

One of Gauss's major publications was *Disquisitiaone Arithmeticae*. In this publication he included a discussion about the Fundamental Theorem of Arithmetic, in which he suggested that its basic principle would also apply to Gaussian integers. He argued that the Gaussian primes will be the basic building blocks of the Gaussian integers, and thus are the basic building blocks of ordinary integers. If this is the case, we need a simple method of identifying which Gaussian integers are prime.

To develop a method of identifying Gaussian primes we need to revisit the method used to show that ordinary primes of the form $m^2 + n^2$ are not Gaussian primes. The first thing we notice is that $m^2 + n^2$ is the square of the modulus. We will define the square of the modulus of a complex number to be norm(z). That is, for any complex number $z = m + ni$, $\text{norm}(z) = m^2 + n^2$.

The norm(z) has an important and interesting property. For any two complex numbers z_1 and z_2, $\text{norm}(z_1 z_2) = \text{norm}(z_1)\text{norm}(z_2)$. It is worth proving this property.

Starting with two complex numbers $z_1 = m + ni$ and $z_2 = p + qi$

$$\begin{aligned}
\text{norm}(z_1)\text{norm}(z_2) &= (m^2 + n^2)(p^2 + q^2) \\
&= (m + ni)(m - ni)(p + qi)(p - qi) \\
&= (m + ni)(p + qi)(m - ni)(p - qi) \\
&= ((mp - nq) + i(mq + np))((mp - nq) - i(mq + np)) \\
&= (mp - nq)^2 + (mq + np)^2
\end{aligned}$$

$$\begin{aligned}
z_1 z_2 &= (mp - nq) + (mq + np)i \\
\text{norm}(z_1 z_2) &= (mp - nq)^2 + (mq + np)^2 \\
\therefore \text{norm}(z_1 z_2) &= \text{norm}(z_1)\text{norm}(z_2)
\end{aligned}$$

We can use this result to show when a Gaussian integer is a Gaussian prime. This is best illustrate by an example.

We will start by showing that $2 + 3i$ is a Gaussian prime.

$$\text{norm}(2 + 3i) = 4 + 9 = 13$$

If $2 + 3i$ has factors which are Gaussian integers, we would expect $z_1 z_2 = 2 + 3i$, where z_1 and z_2 are Gaussian integers.

We know from our previous result that $\text{norm}(z_1 z_2) = \text{norm}(z_1)\text{norm}(z_2)$. This implies that $13 = \text{norm}(z_1)\text{norm}(z_2)$. This requires us to find the factors of 13, but 13 is an ordinary prime. This means that we cannot have *both* $\text{norm}(z_1)$ and $\text{norm}(z_2)$ greater than 1. Thus $2 + 3i$ must be a Gaussian prime.

Let us now consider the Gaussian integer $7 + 3i$.

$$\text{norm}(7 + 3i) = 49 + 9 = 58$$

The factors of 58 are 1, 2, 29 and 58.

$m^2 + n^2 = 2$, when $m = n = \pm 1$

$m^2 + n^2 = 29$, when $m = \pm 5, n = \pm 2$ *or* $m = \pm 2, n = \pm 5$

Using these results:

$$\begin{aligned}
(5 - 2i)(1 + i) &= 5 + 5i - 2i + 2 \\
&= 7 + 3i
\end{aligned}$$

Once again there is a clear relationship between the ordinary primes and the Gaussian primes. We have already shown that an ordinary prime of the form $m^2 + n^2$ is not a Gaussian prime. It would be worth exploring a simple method of identifying this type of prime number.

This can be achieved because of the nature of ordinary prime numbers. With the exception of 2, all ordinary prime numbers are odd numbers. Student Activity 2.3 shows that an even number squared will give an even number and an odd number squared will result in an odd number. So, for $m^2 + n^2$ one of the pair must be odd while the other number must be even. Let us assume that m is even and n is odd.

$$\begin{aligned} m^2 &= (2s)^2 = 4s^2, \text{ where } s \in Z^+ \\ n^2 &= (2r+1)^2 = 4r^2 + 4r + 1, \text{ where } r \in Z^+ \cup \{0\} \\ m^2 + n^2 &= 4s^2 + 4r^2 + 4r + 1 \\ &= 4(s^2 + r^2 + r) + 1 \end{aligned}$$

As both r and s are arbitrary constants, we can let $s^2 + r^2 + r = k$, another arbitrary constant that is a positive integer.

This results suggests that if an ordinary prime can be expressed in the form $4k + 1$, it can also be expressed as $m^2 + n^2$ and, thus, will not be a Gaussian prime.

For $k = 1$ $\quad 4 + 1 = 5 = 1^2 + 2^2$
For $k = 2$ $\quad 8 + 1 = 9$ This is not a prime number.
For $k = 3$ $\quad 12 + 1 = 13 = 3^2 + 2^2$
For $k = 4$ $\quad 16 + 1 = 17 = 4^2 + 1^2$

The Gaussian integers and their primes offer a range of results which display an interplay between Gaussian primes and ordinary primes. This can encourage us to explore all sorts of possibilities, only limited by our imagination. As an example, consider the following activity, which can be used as an investigation activity by students.

STUDENT ACTIVITY 5.4

Ordinary primes of the form $4k + 1$ are not Gaussian primes. Consider the expression $4k + n$, where $n = 0, 1, 2, 3$.

1 Which values of n will generate ordinary prime numbers?
2 Show that an ordinary prime of the form $4k + 3$ will also be a Gaussian prime.

It is of little wonder that Gauss held number theory in such high esteem, particularly when it is extended to include the complex field. We have barely touched the surface of this deep area of wonder. But even this glimpse should encourage use to delve a little deeper.

In the real number field, it has been a challenge since Euclid's time to develop a method of predicting prime numbers. It was believed that there must be a pattern that allows us to develop a formula for the generation of prime numbers. Figure 5.1 shows a set of Gaussian primes that have been plotted onto an Argand diagram using Microsoft Excel. It may make you wonder if there is a pattern for the generation of Gaussian primes, or is it more an image of chaos.

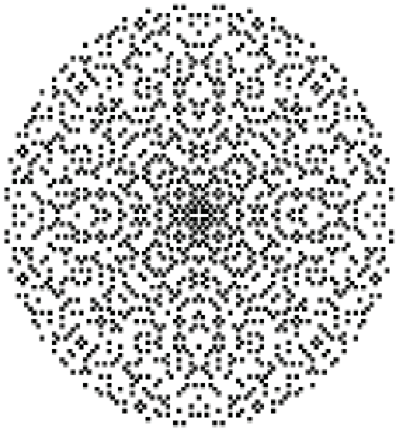


Figure 5.1: An Argand diagram of Gaussian primes

A similar plot appears on the front cover of Bressoud and Wagon (2000).

SUMMARY

- The Fundamental Theorem of Algebra: Every polynomial of degree n with complex coefficients has n roots that can be expressed as complex numbers.
  - When the polynomial has only real coefficients the roots which are complex numbers will appear as conjugate pairs. This is called the conjugate root theorem.
  - Using the polar form of complex numbers allows the use of De Moivre's theorem to solve equations of the form $z^n - a^n = 0$, where a is a real number. Students could be encouraged to use De Moivre's theorem to explore the relationship between the roots of numbers and their position on an Argand diagram.

SUMMARY (Cont.)

- The Fundamental Theorem of Arithmetic: Every integer greater than 1 can be expressed as a product of prime numbers in one, and only one, way.
  - This can be extended to the complex field with the introduction of Gaussian integers.
  - Gaussian primes are Gaussian integers with a modulus greater than 1 which is not the product of Gaussian integers of smaller modulus.
  - Not all ordinary primes are Gaussian primes, for example $5 = (1 - 2i)(1 + 2i)$.
  - To find Gaussian primes it is important that students understand some of the properties of the norm(z), in particular $\text{norm}(z_1 z_2) = \text{norm}(z_1)\text{norm}(z_2)$. This property makes it easier to find Gaussian primes.

References

Bressoud, DM & Wagon, S 2000, *A course in computational number theory,* Springer-Verlag, London.

Websites

http://www-history.mcs.st-andrews.ac.uk/Mathematicians/Gauss.html
This website gives a brief biography of Carl Gauss. It has links to sites that explain many of his mathematical findings.

CHAPTER 6

MATHEMATICIANS CAN READ MAPS

Mappings of complex (numbers and) functions are maps from the complex plane $C = R \times iR$ onto itself. That is, they are a transformation of the plane, which leads to complex analysis and requires consideration of regions of the complex plane. This can be a difficult concept for students to comprehend, but, if treated as an extension of the mapping on the real plane, it can provide for some useful background in senior secondary mathematics.

FUNCTIONS OF A REAL VARIABLE

We have all worked with functions of a real variable. To many of us they are almost second nature. When we see a function such as $f: R \to R, f(x) = x^2$ we can map its image with ease. But we should restate the basics. A function f produces a set of ordered pairs (x ,y) for which the y-value is generated by the rule $y = f(x) = x^2$. The function is operating on a set of x-values that is called the domain of f (Dom f).

When we plot a graph of the function, we plot each ordered pair (x, y). We call y the image of x under the function f. For any function, there are only two aspects we can control: the set of x-values which can be used (its domain) and the rule which creates the image of x.

Essentially the function f maps values of a domain onto its range. The image of a single point is a single point, and the image of all the points from the domain is the range. We can illustrate this with an example (Figure 6.1): $f: \{1, 2, 3\} \to R, f(x) = x^2$.

x	$y = f(x)$
1	1
2	4
3	9

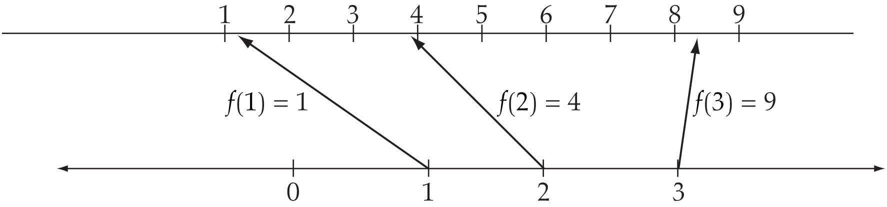


Figure 6.1: Mapping a real function using parallel lines

Instead of using two parallel lines to visualise this mapping, we could rotate one of the lines so that it is perpendicular to the other line. This orientation is called the cartesian plane (Figure 6.2).

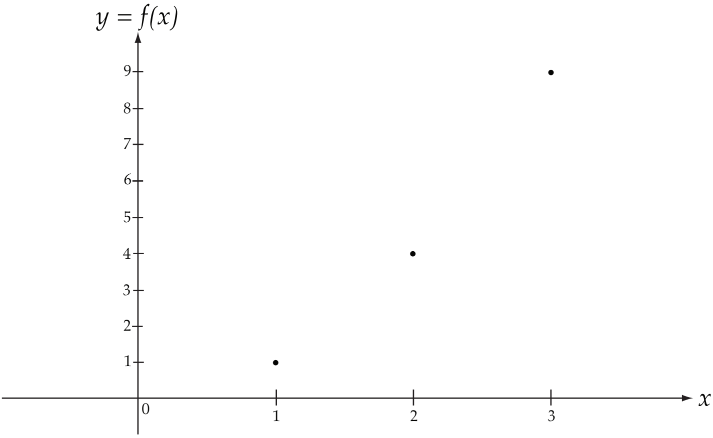


Figure 6.2: Mapping of a real function using perpendicular lines

We could redefine the function, f, to include all positive real values $f: R^+ \rightarrow R$, $f(x) = x^2$. In this case we would expect the image to be a curve because the domain has an infinite set of points which is mapped onto an infinite set of points (Figure 6.3).

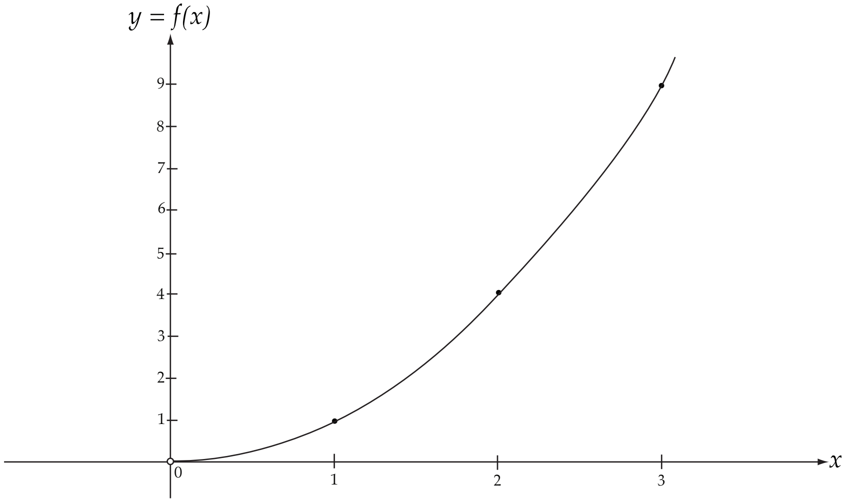


Figure 6.3: Mapping of a real function using perpendicular lines and an infinite set of points

However, zero is not included, as indicated by an open circle at the origin.

FUNCTIONS OF A COMPLEX VARIABLE

We can now turn our attention to functions of a complex variable using the same concepts.

$$f\colon Domain \to C,\ f(z) = z^2$$

The major difference between real and complex functions is the choice of domain. It could be argued that functions of a real variable are a subset of functions of a function of a complex variable. In the case of real variable functions we have restricted the domain to include only the real component of a complex number. The domain of a function of a complex variable will be a subset of the complex numbers, C, and will include real and imaginary components. As a consequence, its range will also be a subset of complex numbers that could contain both real and imaginary components.

When we consider a mapping of these functions we use $w = f(z)$, where w is the image of z under the function f:

$$\begin{aligned} f(z) &= z^2 \\ w &= (x + iy)^2 \\ &= x^2 - y^2 + 2ixy \end{aligned}$$

As we would expect for the function f, the image has a real and an imaginary component. We will need to take this into consideration when we attempt a visual representation of this function. We begin by realising that the image $w = f(z)$ would also have a real and an imaginary component, and can be written as $w = u + iv$, where u and $v \in R$. In senior secondary mathematics it is appropriate to consider some simple complex functions and their effect on subsets of the complex plane.

For $f: C \to C, f(z) = z^2$:

$$w = x^2 - y^2 + 2iy$$

$$\Rightarrow u = x^2 - y^2 \text{ and } v = 2xy$$

To gain an initial impression of this mapping we will use a table of values.

z	x	y	$u = x^2 - y^2$	$v = 2xy$
$1 + i$	1	1	0	2
$2 + i$	2	1	3	4
$3 + i$	3	1	8	6

We will need two complex planes to create this mapping (Figure 6.4).

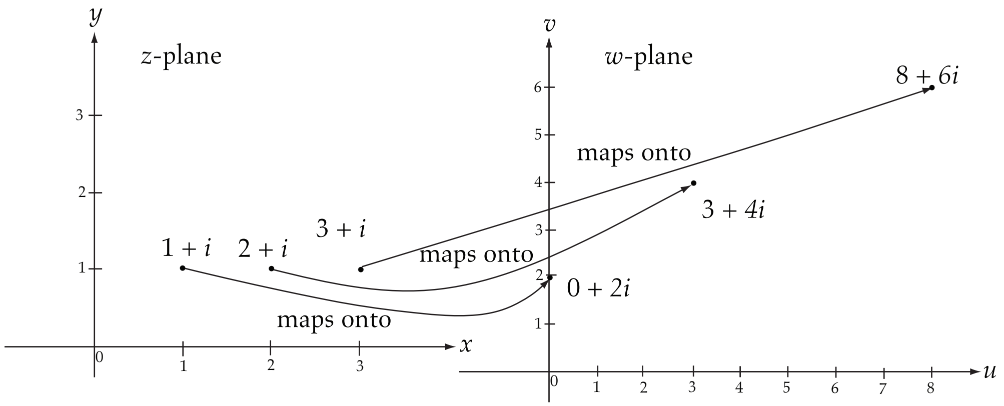


Figure 6.4: Mapping of points on the complex plane

When mapping functions of a complex variable, the mapping is from the z-plane to the w-plane, where the image of the function is given by $w = u(x, y) + v(x, y)i$. Thus u is itself a real function of the variables x, y as is v.

To summarise:

- In the complex number field, a function f maps a complex variable $z = x + iy$ onto a complex image according to a rule specified by $w = f(z)$.

$$f: C \to C, w = f(z)$$

- The function f maps points from its domain in the complex z-plane to points in the complex w-plane, which is a mapping from C to C.

Students should be encouraged to express complex functions in terms of u and v, and thus develop an understanding of the mathematics involved and the mappings that can be generated.

STUDENT ACTIVITY 6.1

For each of the following functions, express $f(z)$ in the form $u + iv$ where u and v are real-value functions of x and y.

a $f: C \to C,\ f(z) = \frac{1}{z}$

b $f: C \to C,\ f(z) = z^3$

c $f: C \to C,\ f(z) = (z + 1)^2$

MAPPING FUNCTIONS OF A COMPLEX VARIABLE

We need a method of exploring the properties of the functions of a complex variable. One way of gaining considerable insight into their properties is to examine images of coordinate lines. A coordinate line is a line that runs parallel to either of the coordinate axes. It can be expressed as either:

$$\{z: z = x + ik,\ \ x, k \in R\}$$

or

$$\{z: z = l + iy,\ \ y, l \in R\}$$

where k and l are *fixed* real constants.

We will use this process to examine the properties of the function of a complex variable $f: C \to C,\ f(z) = z^2 + z$.

The first step is to find u and v in terms of (that is, as functions of) x and y:

$$\begin{aligned} w &= z^2 + z \\ &= (x + iy)^2 + (x + iy) \\ &= x^2 - y^2 + 2ixy + x + iy \\ &= x^2 - y^2 + x + i(2xy + y) \\ \therefore u &= x^2 - y^2 + x \text{ and } v = 2xy + y \end{aligned}$$

We will now find the image of the coordinate line $z = x + ik$ for $k = 1, 2$, and 0.

For $k = 1$ (Figure 6.5):

$$\begin{aligned} u &= x^2 - 1^2 + x \\ &= x^2 + x - 1 \qquad (1) \\ v &= 2x \times 1 + 1 \\ &= 2x + 1 \qquad (2) \end{aligned}$$

From equation (2) $$x = \frac{v-1}{2} \qquad (3)$$

Substituting equation (3) into equation (1):

$$\begin{aligned} u &= \frac{(v-1)^2}{4} + \frac{v-1}{2} - 1 \\ &= \frac{v^2 - 2v + 1 + 2v - 2 - 4}{4} \\ &= \frac{v^2 - 5}{4} \end{aligned}$$

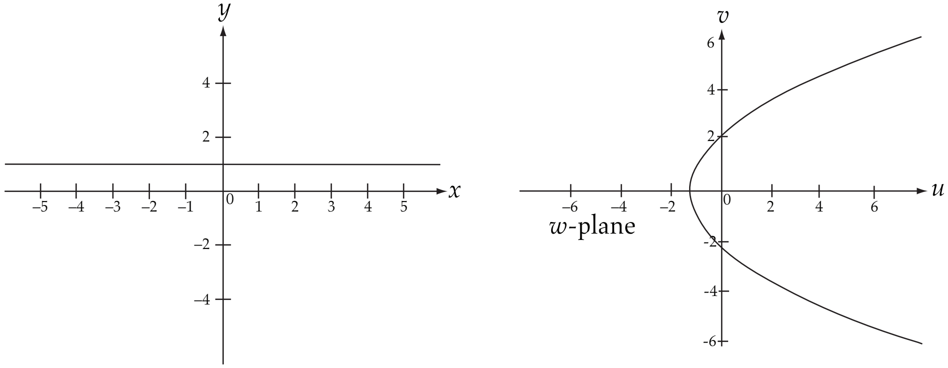


Figure 6.5: Mapping of a line parallel to the *real* axis onto the *w*-plane

For $k = 2$: $$u = x^2 + x - 4$$
$$v = 4x + 2$$
$$x = \frac{v-2}{4}$$
$$\therefore u = \frac{v^2 - 68}{16}$$

For $k = 0$: $$u = x^2 + x$$
$$v = 0$$

For $k = 0$, we have a straight line that runs along the u-axis. However, we need to take into consideration the possible values that u may hold. Given that u is equivalent to the quadratic function $x^2 + x$, we would expect it to hold a minimum value. This minimum value is $\frac{-1}{4}$ which occurs when $x = \frac{-1}{2}$.

The implication of this value is that the image of the line $y = 0$ is the ray which runs along the u-axis where $u \geq \frac{-1}{4}$.

A mapping of the coordinate lines and their images would produce the results shown in Figure 6.6.

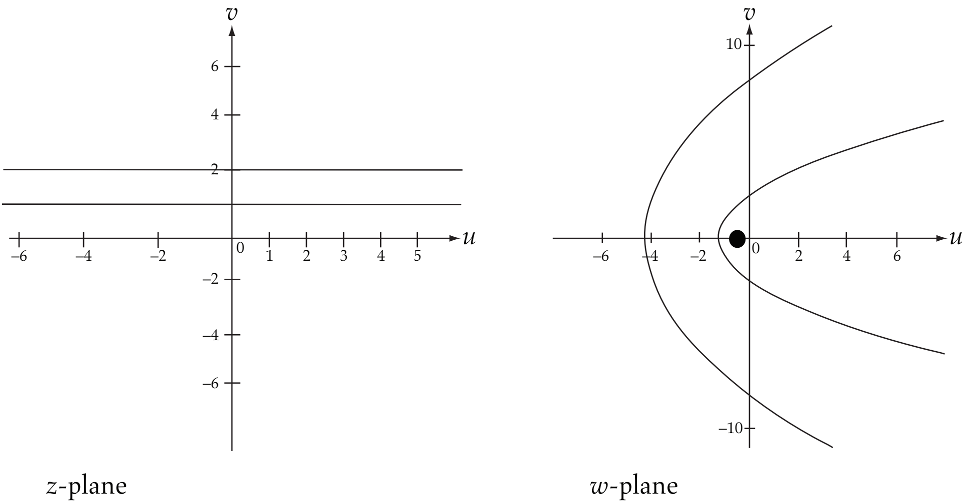


Figure 6.6: Mapping of several lines parallel to the *real* axis onto the *w*-plane

We will now find the image of the coordinate line $z = l + iy$ for $l = 0, 1$ and 2.

For $l = 0$: $u = -y^2$

$$v = y$$

$$u = -v^2$$

For $l = 1$: $u = 2 - y^2$

$$v = 3y$$

$$y = \frac{v}{3}$$

$$\therefore u = 2 - \frac{v^2}{9}$$

For $l = 2$: $u = 6 - y^2$

$$v = 5y$$

$$y = \frac{v}{5}$$

$$\therefore u = 6 - \frac{v^2}{25}$$

When we were finding the images of the line $z = x + ki$, one of the parabolas degenerated into a ray. It would seem reasonable to believe that this may also be the case for the images of the line $z = l + yi$. For this to occur we need to find the value of l that allows $v = 0$.

To find this value we will substitute $z = l + yi$ into $f(z)$.

$$f(l + iy) = l^2 + l - y^2 + i(2ly + y)$$

$$u = l^2 + l - y^2$$

$$v = 2ly + y$$

$$= y(2l + 1)$$

$$v = 0 \text{ when } l = -\tfrac{1}{2}$$

The image of the line $l = -\frac{1}{2}$ occurs when $u = \frac{1}{4} - \frac{1}{2} - y^2$ and $v = 0$.

This implies that we have a ray that runs along the u-axis such that $u \leq \frac{-1}{4}$.

We can now can sketch the graph of the image $w = f(z)$ for several coordinate lines (Figure 6.7).

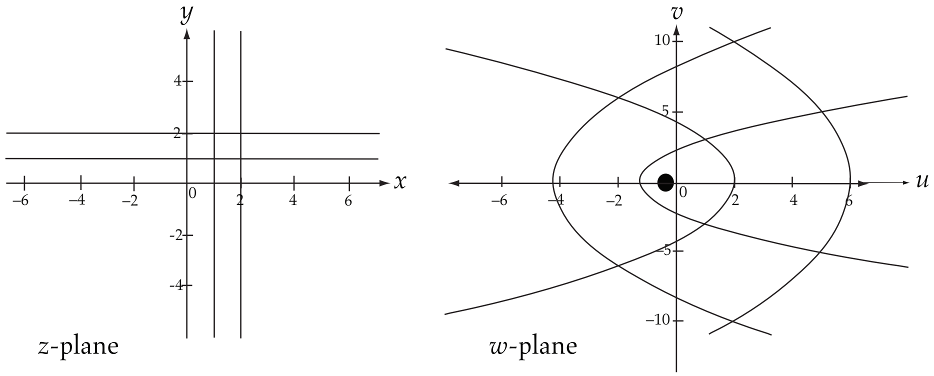


Figure 6.7: Mapping of a line parallel to the *real* and i*maginary* axes onto the *w*-plane

Student Activity 6.2 could be used as an investigation for students to develop mapping on the complex plane.

STUDENT ACTIVITY 6.2

Modern technology allows us a ready ability to map functions of complex variables which can move into representations of Julia and Mandelbrot sets. It is worth visiting the websites http://my.fit.edu/~gabdo/function.html and http://aleph0.clarku.edu/~djoyce/julia/ to discover some of the possibilities. It is also possible to use computer-aided algebra to achieve similar results.

Using a CAS calculator, or otherwise, solve the following.

1 Consider $w = f(z) = z^2$.
   a Find the image of the vertical line $x = l$.
   b Find the image of the horizontal line $y = k$.
2 Consider $w = f(z) = \sqrt{z}$.
   a Find the image of the vertical line $x = l$.
   b Find the image of the horizontal line $y = k$.

THE FUNCTION $f(z) = e^z$

It is worth exploring this function with students because it leads to some interesting results and helps students to realise the relationships between real circular functions and exponential functions.

We will now turn our attention to the mapping of the function:

$$f: C \to C,\ f(z) = e^z$$

Before we can map the image of this function we will need to take a slight detour that will lead to some wonderful discoveries.

We will begin by considering the expansion of e^x:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + \ldots$$

This series converges regardless of the value of x. We will restate this as an expansion for $x = i\theta$:

$$\begin{aligned} e^{i\theta} &= 1 + \frac{i\theta}{1!} + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \ldots + \frac{(i\theta)^n}{n!} + \ldots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^2}{4!} - \frac{\theta^2}{6!} + \ldots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \ldots\right) \end{aligned}$$

The Taylor series expansion of the circular functions sine and cosine are:

$$\cos(\theta) = 1 - \frac{\theta^2}{2!} + \frac{\theta^2}{4!} - \frac{\theta^2}{6!} + \ldots$$

$$\sin(\theta) = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \ldots$$

This suggests that $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. We know from our previous work that $z = \cos(\theta) + i\sin(\theta)$ is the polar form of a complex number with modulus 1. It appears that we now have another representation of the polar form.

The result $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ also leads to some very interesting results that were first given to us by Euler. This is a remarkable result which connects key constants in mathematics: e, π, i and –1.

The first result occurs when we let $\theta = \pi$. This gives us Euler's formula $e^{i\pi} = \cos(\pi) + i\sin(\pi) = -1$.

A second result occurs when we let $\theta = 2\pi$. By substituting this value into the equation $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, we gain the result $e^{2\pi i} = 1$.

We will now return to the mapping of the function $f: C \to C,\ f(z) = e^z$.

We will let $u + iv = f(z)$, where u and v are real.

$$\begin{aligned} u + iv &= e^{x+iy} \\ &= e^x e^{iy} \\ &= e^x(\cos(y) + i\sin(y)) \\ &= e^x\cos(y) + ie^x\sin(y) \end{aligned}$$

Equating real and imaginary components:

$$u = e^x \text{ and } v = e^x \sin(y)$$

By using the coordinate lines:

$$\{z: z = x + ik,\ x, k \in R\}$$

$$\{z: z = l + iy,\ y, l \in R\}$$

We can express relations between u and v for values of k and l:

$$v = u \tan(k)$$

$$u^2 + v^2 = e^{2t}$$

The mapping of the image of $w = f(z)$ for coordinate lines would be concentric circles $u^2 + v^2 = e^{2l}$ that all have the origin as their centre. These circles would be intersected by rays that originate from the origin.

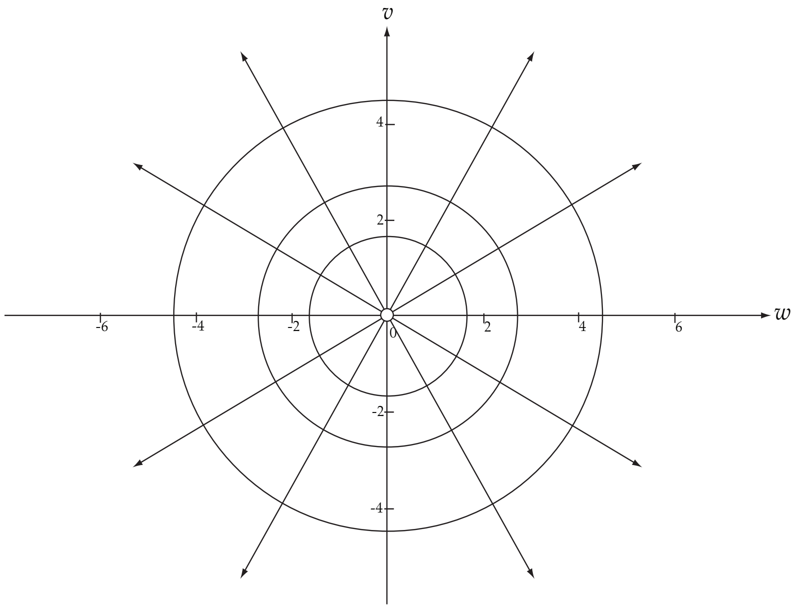


Figure 6.8: The image of $w = f(z)$ for coordinate lines

We know from our study of geometry that any radius of a circle intersects the tangent of a circle at right angles. This suggests that the images of each set of coordinate lines intersect at right angles.

So it seems that the coordinate lines and their images intersect at right angles for the function $f: C \to C,\ f(z) = e^z$. This may also be the case for the function $f: C \to C,\ f(z) = z^2 + z$ and can lead to further investigation.

INTERSECTING TANGENTS

Let us look again at the images of $w = f(z) = z^2 + z$ for the coordinate lines $k = 1$ and $l = 0$.

When $k = 1$, $u = \frac{v^2 - 5}{4}$ and when $l = 0$, $u = -v^2$. These curves intersect at the points $(u, v) = (1, -1)$ and $(-1, -1)$.

For $u = \frac{v^2 - 5}{4}$, $\frac{dv}{du} = \frac{2}{v}$.

When $v = 1$, $\frac{dv}{du} = 2$.

When $v = -1$, $\frac{dv}{du} = -2$.

For $u = -v^2$, $\frac{dv}{du} = \frac{v}{-2}$.

When $v = 1$, $\frac{dv}{du} = \frac{-1}{2}$.

When $v = -1$, $\frac{dv}{du} = \frac{1}{2}$.

We know that if any two gradients are perpendicular, their product will be –1. This is true for both points of intersection. So for the selected values of k and l the images of the tangents at the point where the curves intersect are perpendicular.

Students should be encouraged to explore this property of the mapping of complex functions. Student Activities 5.3 and 5.4 could be used as the basis of an investigation of these properties. These investigations could be supported by the use of an appropriate technology.

STUDENT ACTIVITY 6.3

Show that for all values of l the tangents to the images will be perpendicular to the image generated by the coordinate line given by $k = 1$.

STUDENT ACTIVITY 6.4

Consider the function of the complex variable $f: C \to C$, $f(z) = \frac{1}{z}$.

a Find the equation in terms of u and v of the images of the lines $\{z: z = x + ik,\ x, k \in R\}$ and $\{z: z = l + iy,\ y, l \in R\}$ under f, where k and l are any fixed real number.

b Using appropriate technology, sketch a selection of these curves on an Argand diagram.

c i Find the point of intersection for a pair of curves.

ii Find the angle of the tangents to this pair of curves at their point of intersection.

SETS, CURVES AND REGIONS (WHO ARE OUR NEIGHBOURS?)

Some senior secondary courses include the exploration of subsets of the complex plane such as curves and regions. This discussion is included as background to curves and regions on the complex plane.

The concept of neighbourhood plays an important role in the identification and sketching of complex equations and inequalities on the complex plane. Complex equations and inequalities can be used to represent many different kinds of regions, geometric figures and curves on an Argand diagram.

For a complex number, what is a neighbourhood?

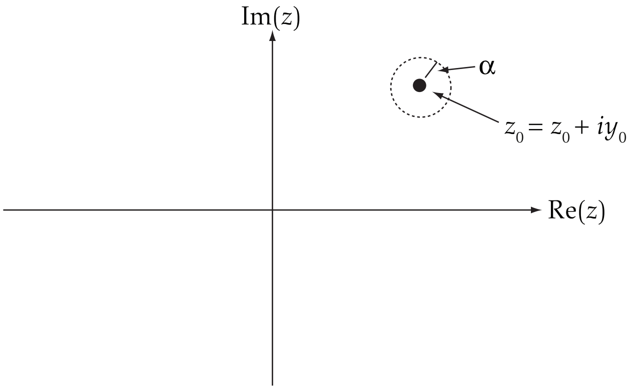


Figure 6.9: Neighbourhood on a complex plane

When we consider a point z_0 on the complex plane, then all the points that are next to this point would be its neighbours. This suggests that the neighbourhood of the point z_0 would be the points that belong to the set S that satisfy the inequality $|z - z_0| < \alpha$, where the value of α is assumed to be very small. This inequality tells us that all the points less than a given distance from the point z_0 are considered to be its neighbours. In geometric terms the neighbourhood of a point z_0 is a disc with radius α (Figure 6.9). Technically, it is not a circle because a circle is the set of points that are an equal distance from a given point and this would be the edge of the disc.

For any neighbourhood, all points of the complex plane have three possible locations. They can be part of the neighbourhood; that is, they are resting on the disc. Such points are referred to as interior points. However, they could be sitting beyond the disc that represents the neighbourhood. Such points are referred to as exterior points.

The boundary of any neighbourhood creates an interesting problem. Consider for a moment the fence line between my house and my neighbour's house (Figure 6.10).

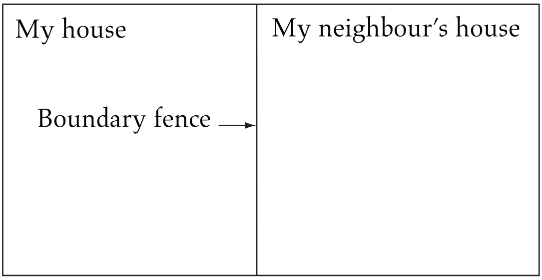


Figure 6.10: The boundary between neighbours

To whom does the fence belong? Does it belong to me or to my neighbour? Does it belong to neither of us? The situation with any such boundary is that it belongs to both of us and to neither of us.

The same situation applies to the boundary between neighbourhoods on the complex plane. A point on the boundary is neither interior nor exterior to S for any neighbourhood. For this reason, when we define any neighbourhood we need to be specific about the inclusion or non-inclusion of the boundary. We should, therefore, define two sets. An open set is a set where all points are interior points; it is a set that does not include the boundary points. A closed set, on the other hand, contains all its boundary points. Any set that consists of only the boundary points is called a curve.

Points in a neighbourhood are considered to be connected if they can be joined by a series of straight lines that are all contained within the set. An open and connected set is called a domain. A neighbourhood is also a domain. We define a *region* to be a domain together with a subset of its boundary. This subset may include, some, part, or all of the boundary points.

Consider the set $|z| \leq 2$ sketched in Figure 6.11. All points on this disc can be connected by a series of straight lines. The set is therefore connected.

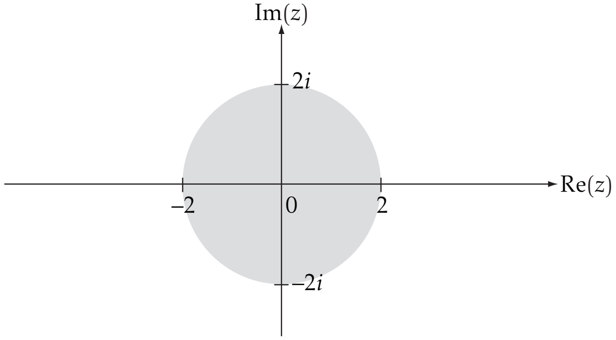


Figure 6.11: A connected set on the complex plane

Students may think these definitions are a bit pedantic and technical, but they should be aware that they need to be careful when they begin to consider regions of a complex plane. So a technical introduction allows them to begin to think more carefully about points on the plane and their relationship with any region. When they sketch regions in the complex plane, they need to ask themselves which points should be included and which should be excluded from the region. The answer to this question is based on careful inspections of the definitions of a given region. Of particular interest is the boundary between the regions they may be sketching.

Now that we have some of the definitions out of the way, we can turn our attention to sketching regions defined by complex equalities or inequalities. We will start with the region $|z - 2 - i| < 3$. We can interpret this inequality geometrically. It is a neighbourhood that has as its centre the complex number $2 + i$. All points are at most 3 units from the centre. This suggests that the inequality represents a disc, and the region does not include the boundary.

To indicate that the boundary is not part of the region we use a dashed line. To indicate when points are part the region we shade the points that are included (Figure 6.12).

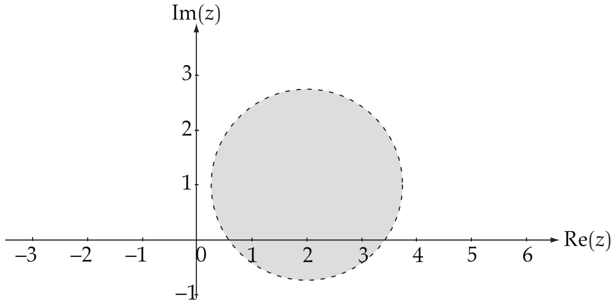


Figure 6.12: The region $|z - 2 - i| < 3$

We can also evaluate $|z-2-i|<3$ algebraically using $z=x+iy$:

$$|z-2-i|<3$$
$$|x+iy-2-i|<3$$
$$|x-2+i(y-1)|<3$$
$$(x-2)^2+(y-1)^2<3^2$$

The equation $(x-2)^2+(y-1)^2=3^2$ is familiar as the cartesian equation of a circle with radius 3 and centre (2, 1). The inequality informs us that we should only consider the values on the interior of the boundary.

We can define interesting curves by considering only the imaginary or real components of a complex equality or inequality.

This can be illustrated by the sketch of $\mathrm{Re}(z^2)>1$ (Figure 6.13).

$$\mathrm{Re}(z^2)>1$$
$$\mathrm{Re}\left((x+iy)^2\right)>1$$
$$\mathrm{Re}\left(x^2-y^2+2ixy\right)>1$$
$$x^2-y^2>1$$

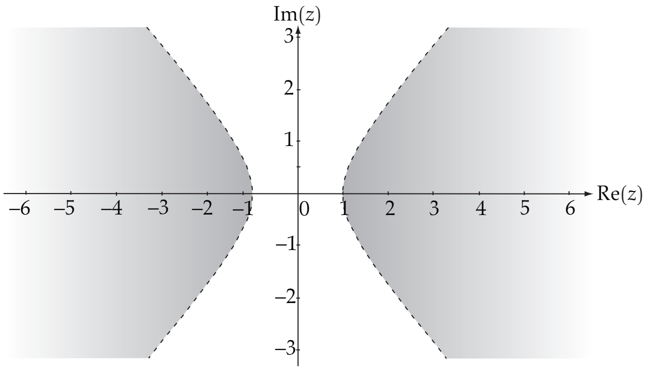


Figure 6.13: The region $\mathrm{Re}(z^2)>1$

It is also possible to use the polar form of a complex number to define either a region or a curve. This can be illustrated by the curve $\mathrm{Arg}(z)=\frac{\pi}{4}$. This is a ray that starts at the origin. It would not include the origin (0, 0) because $\mathrm{Arg}(z)$ is not defined for $z=0$ (Figure 6.14).

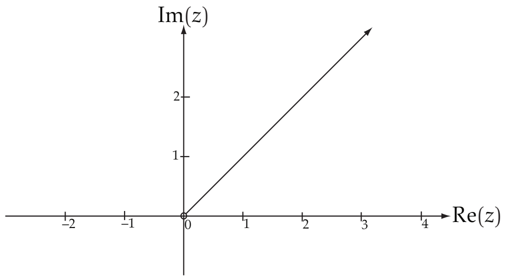


Figure 6.14: The line $\text{Arg}(z) = \frac{\pi}{4}$

We can obtain the related result using algebra:

$$\begin{aligned}
\text{Arg}(z) &= \frac{\pi}{4} \\
\tan^{-1}\left(\frac{y}{x}\right) &= \frac{\pi}{4} \\
\frac{y}{x} &= \tan\left(\frac{\pi}{4}\right) \\
\frac{y}{x} &= 1 \\
y &= x
\end{aligned}$$

This result can very readily create the illusion that $\text{Arg}(z) = \frac{\pi}{4}$ is the straight line $y = x$. But we must be very careful and consider the definition of $\text{Arg}(z)$ and the impact this definition would have on the domain.

This point can be further illustrated by the inequality $\text{Arg}(z-1) \le \frac{\pi}{4}$. Geometrically, we can interpret this to be a ray starting at $x = 0$, but not including this point. The issue now is how to interpret the inequality. We know that $-\pi < \text{Arg}(z) \le \pi$, which means we can only consider points that are in this domain. This means that the boundary that runs along the x-axis should not be included, while we would include the second boundary (Figure 6.15). This example illustrates the importance of the careful consideration of the definitions of any region or curve that may be under consideration.

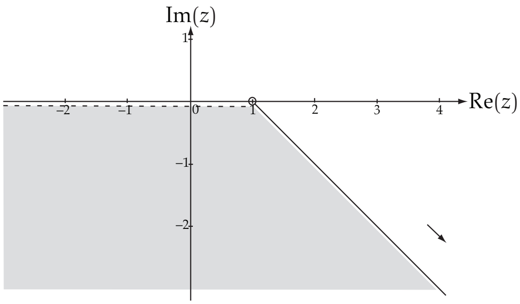


Figure 6.15: The region $\text{Arg}(z-1) \le \frac{\pi}{4}$

Consider one final example, the region defined by $|z-i| \leq |z+1|$. We will begin by considering the boundary. The boundary is all the points which are the same distance from $z=i$ and $z=-1$. This would be the line that is the perpendicular bisector of the line segment that joins these points.

Now we need to be very careful. Do we shade above or below the line? We can use algebra to answer this question.

$$|z-1| \leq |z+1|$$
$$|x+iy-i| \leq |x+iy+1|$$
$$|x+i(y-1)| \leq |x+1+iy|$$
$$x^2+(y-1)^2 \leq (x+1)^2+y^2$$
$$x^2+y^2-2y+1 \leq x^2+2x+1+y^2$$
$$-2y \leq 2x$$
$$y \geq -x$$

The solution confirms that our interpretation of the geometry has given the correct line. It also informs us that the region is all points above the line and the boundary (Figure 6.16).

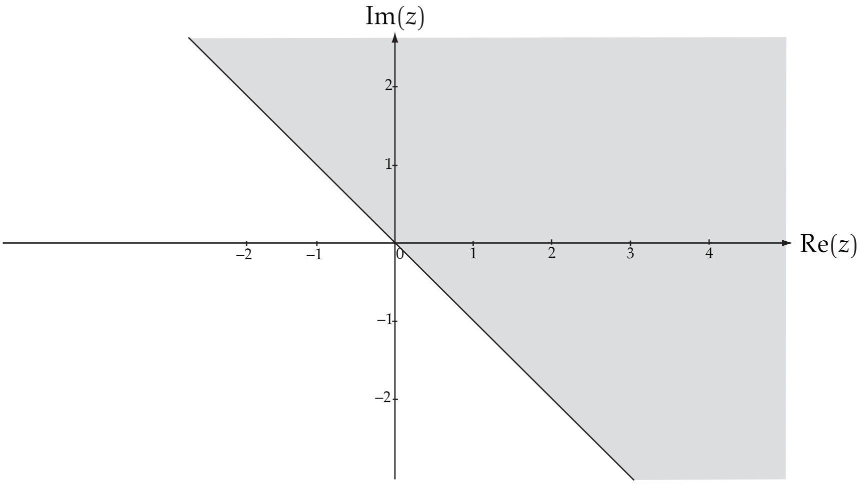


Figure 6.16: The region $|z-i| \leq |z+1|$

There are two aspects of regions of a complex plane students should explore. One is the use of algebra to describe a region (Student Activity 6.5). The second is drawing a region from an algebraic description (Student Activity 6.6).

STUDENT ACTIVITY 6.5

Write an equality or inequality for the following:

a a disc with radius 2 and centre $1 - i$
b all points lying in the 2nd quadrant

STUDENT ACTIVITY 6.6

Draw the locus of points that satisfy the following:

a $3 \leq |z| < 4$

b $|z-1| + |z+i| = 3$

c $|z+i| \leq |z+2|$

d $|z-1| + |z+1| < 3$

Modern technology, in particular graphing packages that use colour, allows us to see the beauty inherent in mapping in the complex field. It is worth undertaking a search on the internet using the key words 'function complex variable' to view some of these mappings.

SUMMARY

- Functions of a complex variable include both imaginary and real components. This means that related mappings will be from one two-dimensional plane onto another two-dimensional plane.
- Students will need to find both the real and imaginary images under the rule of the function. They will need to be careful when considering both the domain and range of the function.
- It is good practice to use coordinate lines to gain an insight into the properties of the mappings of functions of a complex variable.
- The function $f(z) = e^z$ leads to some important relationships, in particular $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, where $\theta = \pi, 2\pi$
- When considering the image of coordinate lines, at intersection points of their images the tangents to the curves are perpendicular.
- Regions and curves can be described on the complex plane using inequalities and equalities involving complex variables. Students need to be careful with inequalities, paying careful attention to the boundaries.

Further reading

Needham, T 1997, *Visual complex analysis*, Oxford University Press, New York.

Websites

http://aleph0.clarku.edu/~djoyce/julia/
This site provides information on Julia and Mandelbrot sets.

http://www.clarku.edu/~djoyce/complex/
This site provides an introduction to complex numbers that includes the mathematics as well as a bit of history.

http://aleph0.clarku.edu/~djoyce/newton/newton.html
Newton's method for finding solutions to equations is explained and leads to some fantastic images when applied to complex functions.

http://my.fit.edu/~gabdo/function.html
This site provides practical experience at plotting functions.

http://faculty.gvsu.edu/fishbacp/dynamics/MandelbrotDemo.htm
This site provides practical experience at plotting Mandelbrot functions.

http://www.pacifict.com/ComplexFunctions.html
This site demonstrates a number of ways to graph a function of a complex variable.

CHAPTER 7

PLOTTING A COURSE

The introduction of vectors to students can provide a number of challenges. Students often find it difficult to visualise vectors, without realising that they intuitively do this every day. The challenge for teachers is to afford students the opportunity to describe this intuition using mathematics. The application of vectors to navigation offers a means of achieving this aim.

The application of vectors to navigation allows students to distinguish between vectors and scalars, as well as understanding many of the key concepts of vectors. It also provides a strong reference point when students attempt to solve problems using vectors.

SEEING THROUGH THE FOG

'Is this fog ever going to lift', wondered Admiral Sir Clowdisley Shovell? 'We need to get our ships and men home safely. It has been a long and successful campaign against the French, and we all deserve to celebrate our victories with family and friends.'

Standing on the quarterdeck of his flagship the *Association*, Admiral Shovell was becoming concerned about the safe return of his fleet. This wasn't the easy trip home he had expected. For 12 days his fleet had been shrouded by fog, its stillness only disturbed by the gentle rhythm of his ship pushing through the ocean.

Sailing in fog created anxiety for all who sailed on wooden ships. All hoped and prayed that the navigators had chosen the correct path or, more importantly, knew exactly where they were.

Sir Clowdisley knew the risks and called all his navigators together and issued two simple instructions: 'Tell me where we are? Is it safe to proceed?'

After several hours of discussion, the navigators were certain that the fleet was east of the Scilly Isles so it was safe to continue north to England.

They were wrong. In fact the ships were further west than the navigators had calculated, and directly in their path lay the Scilly Isles. This miscalculation put the fleet on a collision course.

The flagship was the first to strike the Isles on the night of 22 October 1707. It sank with all hands drowning in an unforgiving sea. It was too late for the rest of the fleet to avoid danger and a further three ships foundered. Only one ship managed to avoid the rocks. The Scilly Isles had become the tombstone of four ships and two thousand men.

The enemy of Sir Clowdisley was not the Isles but his inability to calculate the longitude of the fleet. His was not the only fleet or ship that suffered disaster because of the difficulty of gauging longitude.

Establishing a ship's latitude was relatively easy for any navigator. All you needed to do was know the length of the day and the height of the sun or known stars above the horizon. But for Sir Clowdisley, fog hid both sun and stars. Longitude calculations relied on good seamanship. Navigators relied on 'dead reckoning' to plot their distance east or west. This system meant that navigators needed to find the ship's speed and direction and to factor in ocean currents and wind speed. With a combination of these factors the navigator would establish how far east or west the ship was from the original point.

During the 15th, 16th and 17th centuries European nations relied on their navy captains for exploration and exploitation of the new world. The greatest of these were the best navigators and were accomplished mathematicians who used both the art and science of navigation to avoid disaster and plot a course.

ANCIENT NAVIGATION AND VECTORS

The motion of ships across an ocean offers students a method of visualising vectors. As the ship travels, both its speed and heading (the direction of its journey) need to be taken into account. The representation of a construct that has *both* direction *and* speed will be new to students, although they will be familiar with its application to many everyday events. The challenge for many students is the use of mathematics to describe the familiar.

In mathematics when something has both magnitude and direction it is called a *vector*. When something is given a magnitude but not a direction it is called a *scalar*.

When travelling in a car, we can readily establish our speed (a scalar), but for a navigator speed is only half the story. Travelling at 10 km/hour south-east may allow the safe arrival in port, but 10 km/hour due south may lead to the loss of the ship. Any good navigator knows that when plotting a course it is extremely important to know both the distance and direction of a journey.

We can represent vectors using line segments. As an example we could use a line segment to represent a ship that travels 10 km in a north-east direction, as illustrated in Figure 7.1.

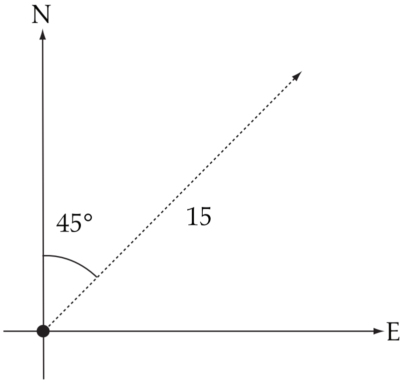


Figure 7.1: Vector diagram for travel in NE direction

Figure 7.2 shows the vector representation of two ships. Ship A has travelled 10 km north-east while ship B has travelled 10 km in an east–north-east direction.

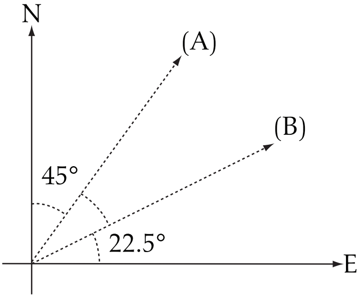


Figure 7.2: Vectors for ships A and B

As vectors are used to give both magnitude and direction, it is essential that we develop a system that clearly distinguishes them from scalars. An obvious visual representation is to use simple directed line segments which give both the start and end points $\overrightarrow{AB}$ (Figure 7.3), but this lacks convenience.

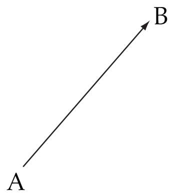


Figure 7.3: A vector or directed line segment

We can name a vector by bolding the letter ***a*** or printing *a* with a tilde (~) underneath the letter $\underset{\sim}{a} = \boldsymbol{a}$, $\underset{\sim}{b} = \boldsymbol{b}$, $\underset{\sim}{c} = \boldsymbol{c}$, $\underset{\sim}{d} = \boldsymbol{d}$ and so on that represents a particular vector.

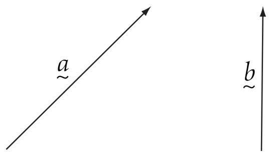


Figure 7.4: Vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$

The magnitude or modulus of a vector is denoted by $|\underset{\sim}{a}|$. It can be visualised as the length of the line segment that represents the vector. It is worth noting that the modulus of a vector or a complex number refers to the same notion.

When considering the journey of Sir Clowdisley's fleet, the distance each ship travels in a given time can be represented by a vector. When both the speed (distance/time) and direction of a ship are considered, this is a vector referred to as *velocity*.

It is interesting to realise that the speed of the fleet only partially contributed to the disaster; of greater significance was the order to turn north too early. This decision bears closer analysis.

The velocity of each ship in the fleet may have been 10 knots north (see Figure 7.5). This notion leads to a definition of the equality of (free) vectors.

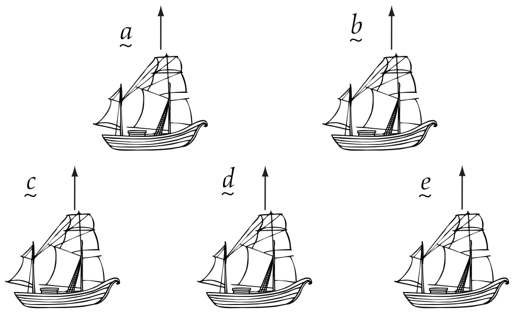


Figure 7.5: Equal velocity of ships

The ships are all travelling in the same direction with the same speed. It seems fair to say that all the ships have the same velocity. That is $\underset{\sim}{a} = \underset{\sim}{b} = \underset{\sim}{c} = \underset{\sim}{d} = \underset{\sim}{e}$.

This notion allows us to understand and define *vector equality*. We can say that all vectors that act in the same direction with the same magnitude are equal.

Expressed in a different way, for any two vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$, if $|\underset{\sim}{a}| = |\underset{\sim}{b}|$ and $\underset{\sim}{a}$ and $\underset{\sim}{b}$ act in the same direction then $\underset{\sim}{a} = \underset{\sim}{b}$.

It becomes clear that knowing the velocity of the fleet will not prevent impending disaster. Further, knowing that the entire fleet had the same velocity is only a partial solution. No doubt Sir Clowdisley was aware that his entire fleet was travelling with equal velocity, but he believed that their position allowed them to be safe.

It would be worth reconsidering the plight of the fleet from the perspective of its position over time. As the fleet sail east into the English Channel, the decision that was required was when to head north. We can represent this journey using vectors (Figure 7.6).

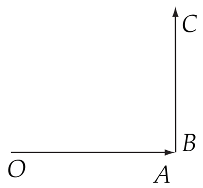


Figure 7.6: Journey represented by vectors

The vector $\overrightarrow{OA}$ represents the fleet's journey from a fixed point O. If the fleet sailed east at 10 km/h for 3 hours, the vector $\overrightarrow{OA}$ has magnitude 30 km and direction east. If the second vector $\overrightarrow{BC}$ represents the ship's journey north for 4 hours, if it maintained its speed of 10 km/h, the vector would have a magnitude of 40 km and direction north.

If the fleet travelled along the two vectors $\overrightarrow{OA}$ and $\overrightarrow{BC}$, it would end up in the same position as if it had travelled directly from the point O to point C (Figure 7.7).

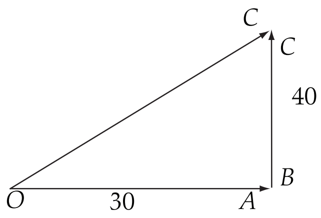


Figure 7.7: A single equivalent vector

It is worth considering this in some detail. It may seem like stating the obvious, but it is worth stating. If the fleet travelled along $\overrightarrow{OA}$ and then $\overrightarrow{BC}$, the consequences are exactly the same as travelling directly to O from C.

We can say this mathematically:

$$\overrightarrow{OA} + \overrightarrow{BC} = \overrightarrow{OC}, \text{ where } \left|\overrightarrow{OC}\right| = 50$$

The direction of $\overrightarrow{OC}$ is given by $\tan^{-1}\left(\frac{40}{30}\right) = 53°7'48''$

The resultant vector $\overrightarrow{OC}$ has magnitude 50 km and direction 53°7′48″ north of west.

This result can be used to help define the addition of vectors.

We will define the addition of two vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ as the single resultant vector that would have the same effect as the combination of $\underset{\sim}{a}$ and $\underset{\sim}{b}$ (Figure 7.8).

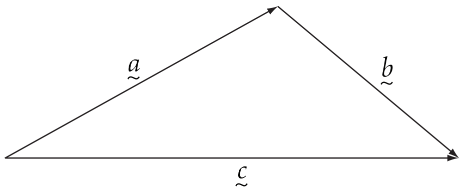


Figure 7.8: The sum of vectors $\underset{\sim}{a} + \underset{\sim}{b} = \underset{\sim}{c}$

That is, $\underset{\sim}{a} + \underset{\sim}{b} = \underset{\sim}{c}$.

Vector $\underset{\sim}{c}$ is the resultant vector of adding the vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$.

When we add vectors it is important to recognise that it is similar to walking the path indicated by the vectors. That is, when we walk along $\underset{\sim}{a}$ in the direction indicated and then along $\underset{\sim}{b}$ in the direction indicated, we will end up in exactly the same position as if we walked the more direct route given by $\underset{\sim}{c}$.

SINKING THE SPANISH ARMADA AND VECTORS

The addition of vectors becomes extremely important when we need to establish the velocity of a ship sailing in an ocean current. This problem proved particularly significant to the defeat of the Spanish Armada.

After several skirmishes with the English Fleet the Spanish attempted to evade the forces of Sir Frances Drake by sailing north into the North Sea. When the Spanish reached the North Sea they intended to sail west until it was safe to travel south back to Spain.

Unfortunately, the Spanish captains were not aware of the Gulf Current that travels northwards along the Irish coast. Let us consider the problem this created for the Spanish captains. As they sailed their ships west they used 'dead reckoning' to establish how far west they had travelled.

The Spanish ships where heavy, slow and hard to steer. Travelling at 7 km/h west they believed that after 10 hours they would have covered 70 km. Using this simple logic, after 6 days the Armada captains believed that it was safe to head south and to the safety of Spain.

Unfortunately they were moving against a current. To understand the consequences, we will simplify the problem by considering two vectors that describe a situation similar to that faced by the Armada. The first is the Gulf Current $\underset{\sim}{C}$ which moves at 5 km/h in a direction E30°N. The second is the Spanish fleet $\underset{\sim}{S}$ which was believed to be travelling on a vector of magnitude 7 km/h with direction W10°N.

The true vector of the ship would be given by adding $\underset{\sim}{S}$ and $\underset{\sim}{C}$.

We will need to use the cosine rule to find the magnitude of the resultant vector:

$$|\underset{\sim}{r}|^2 = 25 + 49 - 70\cos(40°)$$
$$|\underset{\sim}{r}| = 4.51 \text{ km/h}$$

We can then use the sine rule to find the direction in which the fleet was heading:

$$\frac{A}{\sin(a)} = \frac{B}{\sin(b)}$$
$$a = 5 \quad b = r = 4.51$$
$$A = \theta \quad B = 40°$$
$$\frac{r}{\sin(\theta)} = \frac{4.51}{\sin(40)}$$
$$\sin(\theta) = \frac{5\sin(40)}{4.51}$$
$$\theta = \sin^{-1}\left(\frac{5\sin(40)}{4.51}\right)$$
$$= 45.45°$$

Once the ocean current is taken into consideration, the Spanish fleet was travelling at 4.51 km/h on a bearing of W55.45°N because the current has caused a change of 45.45° in their bearing.

It would be worth establishing the true speed at which the fleet was heading west.

This exercise is relatively straightforward using trigonometry (Figure 7.9).

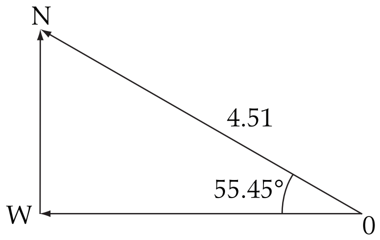


Figure 7.9: Vector representation for the Spanish fleet

$$\left|\overrightarrow{OW}\right| = 4.51\cos(55.45°)$$
$$= 2.56 \text{ km/h}$$

This result indicates that when the fleet has begun to head in a southerly direction it had travelled less than half the distance anticipated by the Spanish captains.

This proved a disaster for the Spanish Armada. There were more ships lost along the Irish coast than to the English fleet.

As you can imagine, the sum of vectors can be difficult when we need to use the cosine and sine rules, particularly if we are required to add more than two vectors.

It would be interesting to re-evaluate the problem faced by Spanish captains in terms of major uses of the compass, north, south, east and west and the addition of vectors, where $\underset{\sim}{N}$ represents a unit vector of magnitude 1 km due north, and similarly for $\underset{\sim}{E}$ and $\underset{\sim}{W}$.

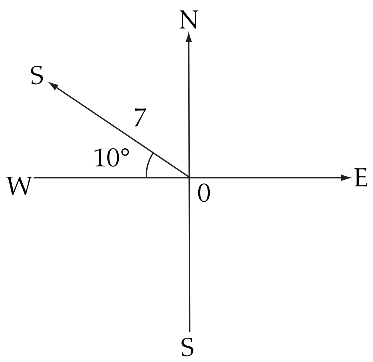


Figure 7.10: Unit vector to represent ships' journey

$$\overrightarrow{OS} = 7\cos(10°)\underset{\sim}{W} + 7\sin(10°)\underset{\sim}{N}$$
$$= 6.89\,\underset{\sim}{W} + 1.22\,\underset{\sim}{N}$$

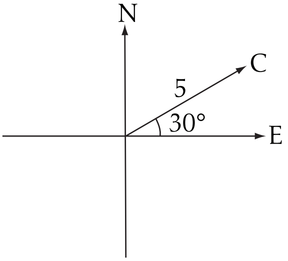


Figure 7.11: Unit vector to represent current

$$\overrightarrow{SC} = 5\cos(30°)\underset{\sim}{E} + 5\sin(30°)\underset{\sim}{N}$$
$$= 4.33\,\underset{\sim}{E} + 2.5\,\underset{\sim}{N}$$
$$\overrightarrow{OS} + \overrightarrow{SC} = 6.89\,\underset{\sim}{W} + 4.33\,\underset{\sim}{E} + 1.22\,\underset{\sim}{N} + 2.5\,\underset{\sim}{N}$$

West and east are opposite directions, so $\overrightarrow{OS} + \overrightarrow{SC}$ can be simplified to become $\overrightarrow{OC} = 2.56\,\underset{\sim}{W} + 3.72\,\underset{\sim}{N}$.

That is, we can break up the vector into separate components. In this case the components are two perpendicular directions north–south and east–west.

Describing a vector as the sum of its components which are mutually perpendicular unit vectors is a powerful means of simplifying vector calculations.

We should extend these unit vectors into three dimensions. While on a two-dimensional ocean it makes sense to use the unit vectors N–S and E–W, we also need to be able to navigate when using aircraft. We can achieve this by using a unit-vector system that uses three unit vectors, $\underset{\sim}{i}$, $\underset{\sim}{j}$ and $\underset{\sim}{k}$, which are oriented as shown in Figure 7.12.

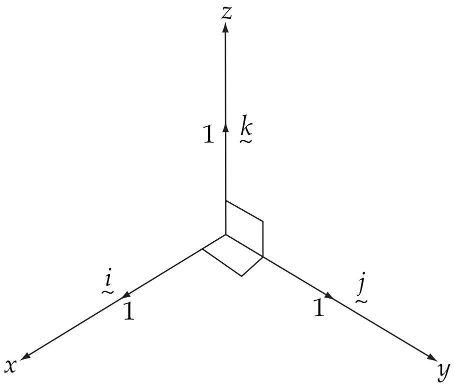


Figure 7.12: Unit vectors in three dimensions

Using this unit-vector system:

$$\overrightarrow{OS} = -6.89\underset{\sim}{i} + 1.22\underset{\sim}{j}$$
$$\overrightarrow{SC} = 4.33\underset{\sim}{i} + 2.5\underset{\sim}{j}$$

When adding vectors in this form we simply collect like terms:

$$\overrightarrow{OS} + \overrightarrow{SC} = (-6.89 + 4.23)\underset{\sim}{i} + (1.22 + 2.5)\underset{\sim}{j}$$
$$= 2.56\underset{\sim}{i} + 3.72\underset{\sim}{j}$$

In general terms, if we have two vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$:

$$\underset{\sim}{a} = x_1\underset{\sim}{i} + y_1\underset{\sim}{j} + z_1\underset{\sim}{k}$$
$$\underset{\sim}{b} = x_2\underset{\sim}{i} + y_2\underset{\sim}{j} + z_2\underset{\sim}{k}$$

then $\underset{\sim}{a} + \underset{\sim}{b} = (x_1 + x_2)\underset{\sim}{i} + (y_1 + y_2)\underset{\sim}{j} + (z_1 + z_2)\underset{\sim}{k}$

The magnitude of a vector in this form will be given by:

$$|\underset{\sim}{a}| = \sqrt{x_1^2 + y_1^2 + z_1^2}$$

We can also consider the subtraction of vectors. In general terms:

$$\underset{\sim}{a} - \underset{\sim}{b} = (x_1 - x_2)\underset{\sim}{i} + (y_1 - y_2)\underset{\sim}{j} + (z_1 - z_2)\underset{\sim}{k}$$

We can also change the magnitude of a vector by multiplying it by a scalar. Once again, consider the following vectors:

$$\underset{\sim}{a} = x_1\underset{\sim}{i} + y_1\underset{\sim}{j} + z_1\underset{\sim}{k}$$
$$\underset{\sim}{d} = c\underset{\sim}{a} = c\left(x_1\underset{\sim}{i} + y_1\underset{\sim}{j} + z_1\underset{\sim}{k}\right) = cx_1\underset{\sim}{i} + cy_1\underset{\sim}{j} + cz_1\underset{\sim}{k}$$

If $c > 0$ then $\underset{\sim}{d}$ is acting in the same direction as $\underset{\sim}{a}$ and has magnitude c times that of $\underset{\sim}{a}$. If $c < 0$, then $\underset{\sim}{d}$ is acting in the opposite direction to $\underset{\sim}{a}$ and has magnitude c times that of $\underset{\sim}{a}$.

We should return to the problem faced by Sir Clowdisley Shovell. He was travelling east through a thick fog into the English Channel. The decision he needed to make was when to turn north. The dangers he faced were twofold:

- If he continued east he was in danger of beaching his fleet on the French coast.
- If he headed north too early he faced the risk of the Scilly Isles.

In a problem similar to that faced by the Spanish Armada, making the right decision relied on correctly gauging the fleet's velocity, and knowing the exact position of the fleet at a given time.

Students need an opportunity to use vectors to solve navigation problems. These can take a variety of forms. Two styles of questions most readily suit this application: the sport of orienteering and the problems faced when navigating an aircraft. The second alternative creates opportunities for students to solve problems in three dimensions. Student Activity 7.1 and Student Activity 7.2 give an indication of the style of question that could be used.

STUDENT ACTIVITY 7.1

In orienteering, runners need to run through a series of checkpoints until they have covered the set course. Good orienteers are both physically fit and good navigators.

For a particular short course an orienteer leaves the starting point heading in a NE direction at a constant speed of 6 km/h. After travelling a distance of 2.5 km he passes the first checkpoint. The next checkpoint is due north at a distance of 1.4 km. On the first leg he is able to maintain a speed of 8 km/h. When he reaches this checkpoint he discovers that the final checkpoint is 3 km to the west over very steep terrain. In this section he can only manage a speed of 2 km/h.

Using $\underset{\sim}{j}$ as north and $\underset{\sim}{i}$ as east, express his velocity and position on each leg in terms of their unit-vector components.

1. How far and at what heading is the orienteer from his starting point?
2. Plot a graph of both his velocity and position for this race.

STUDENT ACTIVITY 7.2

A jet with a cruising speed of 700 km/hour is on a flight from Alice Springs to Darwin, a distance of 1500 km on a bearing directly north. At its cruising altitude, the jet is in an air current of 70 km in a direction S22.5°E.

1. In what direction should the pilot fly the jet to ensure it arrives in Darwin?
2. What is the ground speed of the jet?
3. How long will the journey take?

SUMMARY

- A vector is a mathematical object that can be used to specify both direction and magnitude. A scalar is a mathematical object that only has magnitude.
- Equality of (free) vectors is defined as all vectors that have the same direction and magnitude.
- When vectors are added, both their direction and magnitude must be taken into consideration. Students sometimes make the mistake of only adding vector magnitudes. This only works if the vectors are parallel. The triangle or parallelogram rule for addition of vectors shows how to obtain the sum of two vectors geometrically.
- Using unit vectors $\left(\underset{\sim}{i}, \underset{\sim}{j}, \underset{\sim}{k}\right)$ in three dimensions that are mutually perpendicular (that is *orthogonal*) makes calculation easier.
- Vector operations using unit-vector systems where

  $\underset{\sim}{a} = x_1 \underset{\sim}{i} + y_1 \underset{\sim}{j} + z_1 \underset{\sim}{k}$

  $\underset{\sim}{b} = x_2 \underset{\sim}{i} + y_2 \underset{\sim}{j} + z_2 \underset{\sim}{k}$

  are defined by:
  - addition

    $\underset{\sim}{a} + \underset{\sim}{b} = (x_1 + x_2)\underset{\sim}{i} + (y_1 + y_2)\underset{\sim}{j} + (z_1 + z_2)\underset{\sim}{k}$
  - subtraction

    $\underset{\sim}{a} - \underset{\sim}{b} = (x_1 - x_2)\underset{\sim}{i} + (y_1 - y_2)\underset{\sim}{j} + (z_1 - z_2)\underset{\sim}{k}$
  - multiplication by a scalar

    $c\underset{\sim}{a} = cx_1 \underset{\sim}{i} + cy_1 \underset{\sim}{j} + cz_1 \underset{\sim}{k}$

Further reading

Sobel, D 1995, *Longitude*, Walker, New York.
Wason, D 2003, *Battlefield detectives*, Granada Media, London.

Websites

There are a number of websites that offer 'vector calculators'. These online calculators can simplify some vector calculations. Two examples of these websites are listed.
http://comp.uark.edu/%7Ejgeabana/java/VectorCalc.html
http://hyperphysics.phy-astr.gsu.edu/hbase/vect.html

CHAPTER 8

SAILING AGAINST THE WIND

A natural approach to the introduction of vector *products* is to apply them practically in several useful contexts. These contexts should be developed to allow students to *visualise* the application of vector products. Vector products have powerful applications in physics; for example, the vector *scalar* product can be used measure work, while the vector *cross* product can be used to measure torque. They can also be used in geometry. An interesting application is to use them in combination to find the volume of a parallelepiped. As a simple introduction to the vector scalar product it is useful to begin with a similar context to the previous chapter—navigation of sailing ships.

TACKING INTO THE WIND

It wasn't so long ago that the power that drove commerce and trade around the globe was the wind. People and goods travelled the oceans in tall-masted wooden ships.

It is worth stepping back and reflecting on the courage and skill of those who sailed on wooden ships. In the previous chapter we learned to appreciate the risks and skills of those who use their mathematical skills to navigate their ships from port to port.

For present-day sailors who rely on the power of fossil fuels, navigation systems can plot a course directly to a location. Sailors can check their progress using global positioning systems. This is true even when they are required to travel into the wind.

Sailing into the wind required great skill from the crew and navigators of tall ships. It required a navigator to plot a course that zigged and zagged towards the desired point. This process is called tacking.

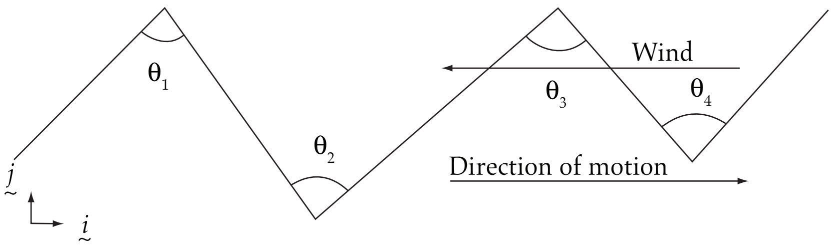


Figure 8.1: Tacking into the wind

A problem that needs to be considered is the angle θ_n, the angle through which the ship turns between successive tacks into the wind.

Let's consider one angle turned through in a single tack.

At the point of the tack the ship's position could be given by $\overrightarrow{OP_1} = 8\underset{\sim}{i} + 2\underset{\sim}{j}$.

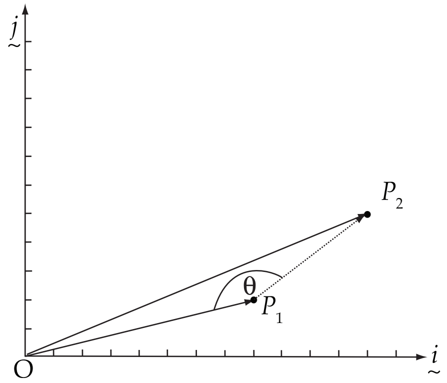


Figure 8.2: Graphical representation of a tack

Its position at the next tack is $\overrightarrow{OP_2} = 12\underset{\sim}{i} + 5\underset{\sim}{j}$.

The vector $\overrightarrow{P_1P_2} = \overrightarrow{OP_2} - \overrightarrow{OP_1} = 4\underset{\sim}{i} + 3\underset{\sim}{j}$ (Figure 8.2).

It is possible to use the cosine rule to find the angle θ.

The first step is to find the angle θ by reconstructing the triangle with the magnitude of each vector as the length of each side of the triangle (Figure 8.3).

The magnitude of a vector $\underset{\sim}{a} = x\underset{\sim}{i} + y\underset{\sim}{j} + z\underset{\sim}{k}$ is given by:

$$\left|\underset{\sim}{a}\right| = \sqrt{x^2 + y^2 + z^2}$$

$$\left|\overrightarrow{OP_2}\right| = \sqrt{144 + 25}$$
$$= \sqrt{169}$$
$$= 13$$

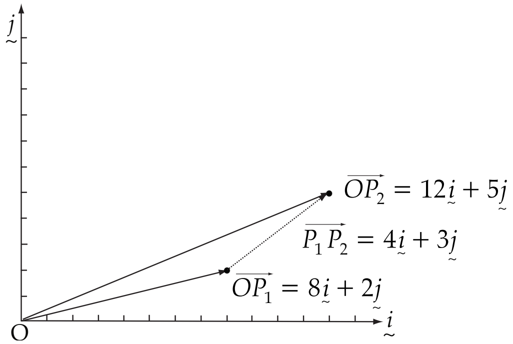


Figure 8.3: Triangle formed by a tack

Extracting the triangle (Figure 8.4):

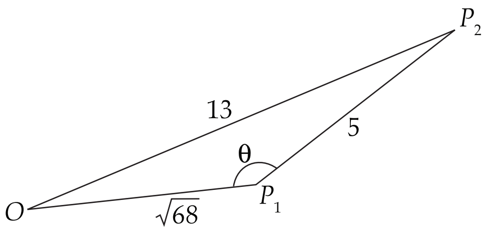


Figure 8.4: Extracted triangle for a tack

Using the cosine rule:

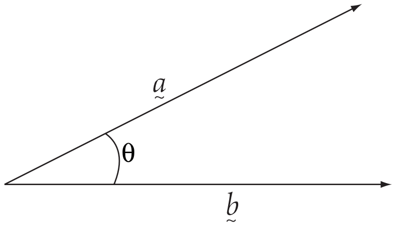
$$a^2 = b^2 + c^2 - 2bc\cos(\theta)$$
$$169 = 68 + 25 - 2 \times \sqrt{68} \times 5\cos(\theta)$$
$$\cos(\theta) = \frac{-76}{10\sqrt{68}}$$
$$\theta = \cos^{-1}\left(\frac{-19}{5\sqrt{17}}\right)$$

There is another method of solving this problem which uses a form of vector multiplication. There are two ways in which vectors can be multiplied. The method that allows for a simpler method of finding the angle between successive tacks of tall ship is called the *scalar* or dot product of vectors.

It should be emphasised to students that this is not the same as the scalar multiple of a vector.

THE SCALAR OR DOT PRODUCT OF TWO VECTORS

Figure 8.5 shows the definition of the scalar or dot product of two vectors, $\underset{\sim}{a}$ and $\underset{\sim}{b}$.



$$\underset{\sim}{a}.\underset{\sim}{b} = |\underset{\sim}{a}||\underset{\sim}{b}|\cos(\theta)$$

Figure 8.5: Definition of scalar or dot product

To improve our appreciation of the usefulness of the dot product it is worth considering the concept of 'work' as defined by physicists. For physicists, work done is given by the magnitude of a force applied to an object multiplied by the distance the object has moved in the direction of the motion.

Consider as an example the work done by a horse as it tows a barge along a canal (Figure 8.6).

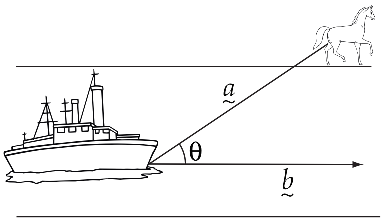


Figure 8.6: Vector representation of barge being towed

$\underset{\sim}{a}$ is the force applied to the barge while $\underset{\sim}{b}$ is the displacement of the barge. The force applied in the direction of the motion would be $|\underset{\sim}{a}|\cos(\theta)$.

The work done by the horse in moving the barge would be $|\underset{\sim}{a}||\underset{\sim}{b}|\cos(\theta)$.

To use the least amount of force to generate the same amount of work requires the angle θ to be equal to zero. This is not possible in this case, unless the horse can swim and pull the barge at the same time. However, if we increase the length of the tow rope, the horse requires less force to achieve the same amount of work.

It is also worth noting that if the horse attempted to tow the barge by walking at an angle of 90° to the direction of the path no work would be done. The dot product deals with this notion because $\cos(90^\circ)$ is zero.

It is important to realise that when we use the dot product that $|\underset{\sim}{a}|$, $|\underset{\sim}{b}|$ and $\cos(\theta)$ are all numbers or scalar quantities. This means that $\underset{\sim}{a}.\underset{\sim}{b}$ is not a vector but a scalar quantity. This is why it is called the *scalar product*.

The ability to use the scalar product is greatly simplified when we use $\underset{\sim}{i}$, $\underset{\sim}{j}$ and $\underset{\sim}{k}$ components to represent vectors.

Since $\underset{\sim}{i}$, $\underset{\sim}{j}$ and $\underset{\sim}{k}$ are mutually perpendicular and 1 unit in magnitude, the scalar product for each combination would be, by definition:

$$\underset{\sim}{i}.\underset{\sim}{i} = \underset{\sim}{j}.\underset{\sim}{j} = \underset{\sim}{k}.\underset{\sim}{k} = 1$$

and

$$\underset{\sim}{i}.\underset{\sim}{j} = \underset{\sim}{i}.\underset{\sim}{k} = \underset{\sim}{j}.\underset{\sim}{k} = 0$$

We can use this information to define scalar multiplication using $\underset{\sim}{i}$, $\underset{\sim}{j}$ and $\underset{\sim}{k}$ components of a vector:

if $$\underset{\sim}{a} = x_1\underset{\sim}{i} + y_1\underset{\sim}{j} + z_1\underset{\sim}{k}$$

and $$\underset{\sim}{b} = x_2\underset{\sim}{i} + y_2\underset{\sim}{j} + z_2\underset{\sim}{k}$$

then $$\underset{\sim}{a}.\underset{\sim}{b} = x_1x_2 + y_1y_2 + z_1z_2$$

Let's return to the angle θ through which a tall ship travels on successive tacks (Figure 8.7).

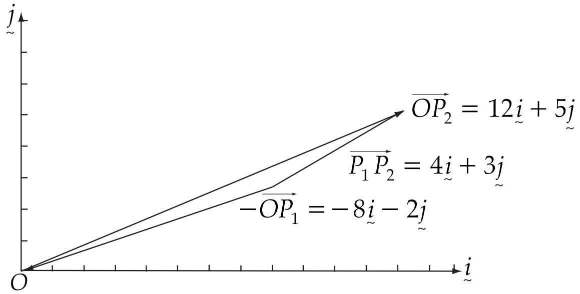


Figure 8.7: Successive tacks for a tall ship, adjusted for dot product

$$
\begin{aligned}
\text{Let} \qquad \underset{\sim}{a} &= -\overrightarrow{OP_1} = -8\underset{\sim}{i} - 2\underset{\sim}{j} \\
\underset{\sim}{b} &= \overrightarrow{P_1P_2} = 4\underset{\sim}{i} + 3\underset{\sim}{j} \\
\underset{\sim}{a}.\underset{\sim}{b} &= |\underset{\sim}{a}||\underset{\sim}{b}|\cos(\theta) \\
-32 - 6 &= 2\sqrt{17} \times 5\cos(\theta) \\
\cos(\theta) &= \frac{-38}{10\sqrt{17}} = \frac{-19}{5\sqrt{17}}
\end{aligned}
$$

PERPENDICULAR VECTORS

The scalar product is a very useful method of finding whether two vectors are perpendicular (orthogonal).

Consider the scalar product $\underset{\sim}{a}.\underset{\sim}{b} = |\underset{\sim}{a}||\underset{\sim}{b}|\cos(\theta)$

If $|\underset{\sim}{a}| \neq 0$ and $|\underset{\sim}{b}| \neq 0$ then $\underset{\sim}{a}.\underset{\sim}{b}$ will equal zero if $\cos(\theta) = 0$. This only occurs when $\theta = 90°$

So if $|\underset{\sim}{a}| \neq 0$ and $|\underset{\sim}{b}| \neq 0$ and $\underset{\sim}{a}.\underset{\sim}{b} = 0$, we know that this can only occur if $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are orthogonal.

This can be illustrated by an example:

$$
\begin{aligned}
\underset{\sim}{a} &= 3\underset{\sim}{i} - 5\underset{\sim}{j} - 2\underset{\sim}{k} \\
\underset{\sim}{b} &= 2\underset{\sim}{i} + 2\underset{\sim}{j} - 2\underset{\sim}{k} \\
\underset{\sim}{a}.\underset{\sim}{b} &= 6 - 10 + 4 = 0
\end{aligned}
$$

This means that the vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are orthogonal.

This gives two rationales for the introduction of the scalar product. It can be used to describe the physics notion of work and to establish that the vectors are perpendicular (orthogonal). Another useful application for the scalar product is to find the projections of vectors in a given (arbitrary) direction. The direction cosines of vectors is one method of achieving the latter application.

DIRECTION COSINES OF A VECTOR

Another interesting application of the scalar product is the ability to find the angle a vector makes with each of the cartesian axes in three dimensions (Figure 8.8).

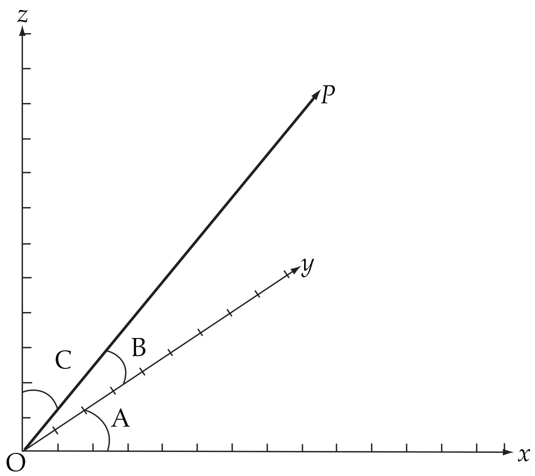


Figure 8.8: A vector $\overrightarrow{OP}$ in three-dimensional space

The vector $\overrightarrow{OP} = a\underset{\sim}{i} + b\underset{\sim}{j} + c\underset{\sim}{k}$

Using the scalar product:

$$\begin{aligned} \underset{\sim}{P}.\underset{\sim}{i} &= |\underset{\sim}{P}||\underset{\sim}{i}|\cos(A) \\ \Rightarrow a &= |\underset{\sim}{P}|\cos(A) \\ \therefore \cos(A) &= \frac{a}{|\underset{\sim}{P}|} \end{aligned}$$

In a similar way:

$$\cos(B) = \frac{b}{|\underset{\sim}{P}|}$$

and
$$\cos(C) = \frac{c}{|\underset{\sim}{P}|}$$

We also know that $|\underset{\sim}{P}| = \sqrt{a^2 + b^2 + c^2}$.

Using the cosines of the angles vector $\underset{\sim}{P}$ makes with each of the axes—called the direction cosines—we can write down a vector in terms of its direction only:

$$\underset{\sim}{r} = \cos(A)\underset{\sim}{i} + \cos(B)\underset{\sim}{j} + \cos(C)\underset{\sim}{k}$$

This can also be written as:

$$\underset{\sim}{r} = \frac{a}{|\underset{\sim}{P}|}\underset{\sim}{i} + \frac{b}{|\underset{\sim}{P}|}\underset{\sim}{j} + \frac{c}{|\underset{\sim}{P}|}\underset{\sim}{k}$$
$$= \frac{1}{|\underset{\sim}{P}|}\left(a\underset{\sim}{i} + b\underset{\sim}{j} + c\underset{\sim}{k}\right)$$

The magnitude of $\underset{\sim}{r}$ would be:

$$|\underset{\sim}{r}| = \frac{\sqrt{a^2 + b^2 + c^2}}{|\underset{\sim}{P}|}$$
$$= \frac{\sqrt{a^2 + b^2 + c^2}}{\sqrt{a^2 + b^2 + c^2}} = 1$$

This means that $\underset{\sim}{r}$ is a unit vector in the direction of $\underset{\sim}{P}$. We will define this form of unit vector as $\hat{\underset{\sim}{P}}$ (P hat), where:

$$\hat{\underset{\sim}{P}} = \cos(A)\underset{\sim}{i} + \cos(B)\underset{\sim}{j} + \cos(C)\underset{\sim}{k}$$
$$= \frac{a\underset{\sim}{i} + b\underset{\sim}{j} + c\underset{\sim}{k}}{|\underset{\sim}{P}|}$$
$$= \frac{\underset{\sim}{P}}{|\underset{\sim}{P}|}$$

$\hat{\underset{\sim}{P}}$ is referred to as the unit vector in the direction of $\underset{\sim}{P}$.

PARALLEL AND PERPENDICULAR COMPONENTS FOR $\underset{\sim}{P}$

It is possible to use the scalar product to find components of a vector $\underset{\sim}{r}$ that run perpendicular to and parallel to a second vector $\underset{\sim}{p}$.

This can be represented on a diagram.

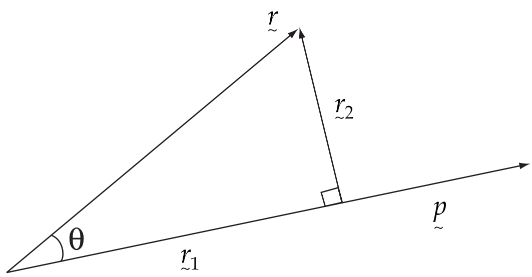


Figure 8.9: Parallel and perpendicular components of a vector

The magnitude of $|\underset{\sim}{r}_1|$ is given by $|\underset{\sim}{r}|\cos(\theta)$.

We can use $\hat{\underset{\sim}{P}}$ to find $|\underset{\sim}{r}|\cos(\theta)$:

$$\hat{\underset{\sim}{P}}.\underset{\sim}{r} = |\underset{\sim}{P}||\underset{\sim}{r}|\cos(\theta)$$

This means that $|\underset{\sim}{r}| = |\hat{\underset{\sim}{P}}||\underset{\sim}{r}|\cos(\theta)$, given that $\hat{\underset{\sim}{P}}$ is a unit vector in the direction of $\underset{\sim}{P}$.

$$\begin{aligned}\underset{\sim}{r}_1 &= |\underset{\sim}{r}_1|\hat{\underset{\sim}{P}} \\ &= |\hat{\underset{\sim}{P}}||\underset{\sim}{r}|\cos(\theta)\hat{\underset{\sim}{P}} \\ &= (\hat{\underset{\sim}{P}}.\underset{\sim}{r})\hat{\underset{\sim}{P}}\end{aligned}$$

From Figure 8.9 it can be seen that:

$$\begin{aligned}\underset{\sim}{r}_2 &= \underset{\sim}{r}_1 - \underset{\sim}{r} \\ &= (\hat{\underset{\sim}{P}}.\underset{\sim}{r})\hat{\underset{\sim}{P}} - \underset{\sim}{r}\end{aligned}$$

Thus, any vector can be expressed as the sum of two vectors, one parallel and one perpendicular to any other given vector.

THE CROSS PRODUCT OF TWO VECTORS

When using the scalar product, the result of multiplying two vectors is not a vector. It would be useful to be able to multiply two vectors so that the result is a *vector*.

If we consider two vectors $\underset{\sim}{a}, \underset{\sim}{b}$ we need to find a way of combining them with a single direction for the resulting vector.

This cannot be achieved if we restrict ourselves to two dimensions.

We know that two vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ will be on a single plane (Figure 8.10).

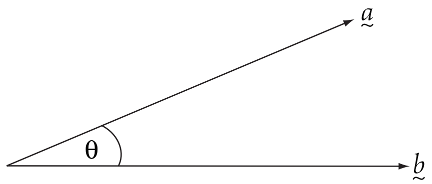


Figure 8.10: Two vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$

To assign a direction to the resultant vector we will use a vector that is perpendicular to the plane (Figure 8.11).

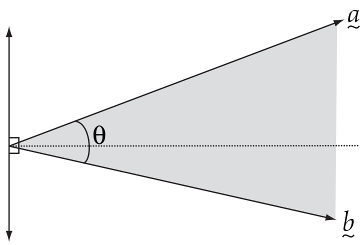


Figure 8.11: Two vectors $\utilde{a}$ and $\utilde{b}$ with a vector perpendicular to the plane

Using this idea it is possible to assign two directions that are perpendicular to the plane.

The direction that will be assigned to the resultant vector will use the convention used to assign direction to cartesian coordinates, and the component vectors $\utilde{i}$, $\utilde{j}$ and $\utilde{k}$.

A simple method of visualising this orientation of vectors is to use the right hand as a reference.

Using the right-hand rule, the thumb represents the $\utilde{i}$ component, the index finger the $\utilde{j}$ component and the middle finger the $\utilde{k}$ component.

We can now define the vector cross product as $\utilde{a} \times \utilde{b} = |\utilde{a}||\utilde{b}|\sin(\theta)\hat{\utilde{n}}$.

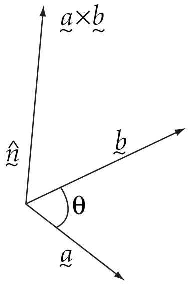


Figure 8.12: Visual representation of the vector cross product

$\hat{\utilde{n}}$ is a unit vector perpendicular to the plane in which both $\utilde{a}$ and $\utilde{b}$ lie. We need to realise that given two vectors $\utilde{a}, \utilde{b}$ the set of all combinations $\utilde{u} = m\utilde{a} + n\utilde{b}$ defines a plane. The unit vector $\hat{\utilde{n}}$ will be perpendicular to this plane in both the 'up' and 'down' directions, so we need to define its direction. We use the *right-hand rule*, in a similar way to its previous use, to define the direction of $\hat{\utilde{n}}$.

VECTOR COMPONENTS AND THE CROSS PRODUCT

Given that $\sin(0°) = 0$ and $\sin(90°) = 1$ and the vectors $\underset{\sim}{i}$, $\underset{\sim}{j}$ and $\underset{\sim}{k}$ are unit vectors that are mutually perpendicular:

$$\underset{\sim}{i} \times \underset{\sim}{i} = \underset{\sim}{j} \times \underset{\sim}{j} = \underset{\sim}{k} \times \underset{\sim}{k} = 0$$

$$\underset{\sim}{i} \times \underset{\sim}{j} = \underset{\sim}{k},\ \underset{\sim}{j} \times \underset{\sim}{k} = \underset{\sim}{i} \text{ and } \underset{\sim}{k} \times \underset{\sim}{i} = \underset{\sim}{j}$$

Also keeping in mind the right-hand rule:

$$\underset{\sim}{j} \times \underset{\sim}{i} = -\underset{\sim}{k},\ \underset{\sim}{k} \times \underset{\sim}{j} = -\underset{\sim}{i} \text{ and } \underset{\sim}{i} \times \underset{\sim}{k} = -\underset{\sim}{j}$$

Consider the two vectors $\underset{\sim}{a} = x_1\underset{\sim}{i} + y_1\underset{\sim}{j} + z_1\underset{\sim}{k}$ and $\underset{\sim}{b} = x_2\underset{\sim}{i} + y_2\underset{\sim}{j} + z_2\underset{\sim}{k}$.

$$\begin{aligned} \underset{\sim}{a} \times \underset{\sim}{b} = {} & x_1x_2(\underset{\sim}{i} \times \underset{\sim}{i}) + x_1y_2(\underset{\sim}{i} \times \underset{\sim}{j}) + x_1z_2(\underset{\sim}{i} \times \underset{\sim}{k}) \\ & + y_1x_2(\underset{\sim}{j} \times \underset{\sim}{i}) + y_1y_2(\underset{\sim}{j} \times \underset{\sim}{j}) + y_1x_2(\underset{\sim}{j} \times \underset{\sim}{k}) \\ & + x_1z_1(\underset{\sim}{k} \times \underset{\sim}{i}) + y_2z_1(\underset{\sim}{k} \times \underset{\sim}{j}) + z_1z_2(\underset{\sim}{k} \times \underset{\sim}{k}) \\ = {} & 0 + x_1y_2\underset{\sim}{k} + x_1z_2(-\underset{\sim}{j}) + y_1x_2(-\underset{\sim}{k}) + 0 + y_1z_2\underset{\sim}{i} + x_2z_1\underset{\sim}{j} + y_2z_1(-\underset{\sim}{i}) + 0 \\ = {} & (y_1z_2 - y_2z_1)\underset{\sim}{i} + (z_1x_2 - x_1z_2)\underset{\sim}{j} + (x_1y_2 - x_2y_1)\underset{\sim}{k} \end{aligned}$$

The following discussion shows some applications of the vector cross product.

MEASURING TORQUE

An application of the cross product is the measurement of *torque*.

Torque measures the *turning effect* of a force $\underset{\sim}{F}$ about an axis. Consider a force $\underset{\sim}{F}$ acting at an angle θ to $\underset{\sim}{r}$. Then, as shown in Figure 8.13

$$\underset{\sim}{T} = \underset{\sim}{r} \times \underset{\sim}{F} = |\underset{\sim}{r}||\underset{\sim}{F}|\sin(\theta)\underset{\sim}{n}$$

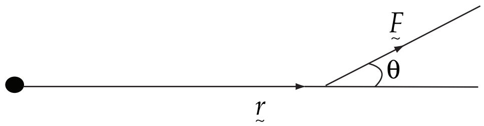


Figure 8.13: Measuring torque

MULTIPLYING VECTORS THREE AT A TIME

Consider three vectors $\underset{\sim}{a}$, $\underset{\sim}{b}$ and $\underset{\sim}{c}$. Which of the following combinations are possible?

$$(\underset{\sim}{a}.\underset{\sim}{b}).\underset{\sim}{c} \tag{1}$$

$$(\underset{\sim}{a}.\underset{\sim}{b})\times\underset{\sim}{c} \tag{2}$$

$$(\underset{\sim}{a}\times\underset{\sim}{b})\times\underset{\sim}{c} \tag{3}$$

$$(\underset{\sim}{a}\times\underset{\sim}{b}).\underset{\sim}{c} \tag{4}$$

Neither (1) nor (2) is possible because $\underset{\sim}{a}.\underset{\sim}{b}$ gives a scalar result.

Expression (3) will be possible because $\underset{\sim}{a}\times\underset{\sim}{b}$ gives a vector result. This will also be true for (4). This last case deserves greater exploration because it can be used to find the volume of a slant-sided box, or parallelepiped.

VOLUME OF A PARALLELEPIPED

We will start by considering the simple rectangular prism formed by these vectors (Figure 8.14):

$$\underset{\sim}{a} = 3\underset{\sim}{i}$$
$$\underset{\sim}{b} = 4\underset{\sim}{j}$$
$$\underset{\sim}{c} = 5\underset{\sim}{k}$$

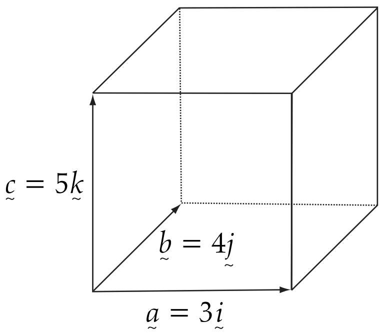


Figure 8.14: Cube formed by $3\underset{\sim}{i}$, $4\underset{\sim}{j}$ and $5\underset{\sim}{k}$

If we consider the magnitude of these three vectors, then the volume of the cube is simply $3\times4\times5 = 60$ cubic units.

Now
$$\begin{aligned} \underset{\sim}{a} \times \underset{\sim}{b} &= 3\underset{\sim}{i} \times 4\underset{\sim}{j} \\ &= 12\underset{\sim}{k} \\ (\underset{\sim}{a} \times \underset{\sim}{b}).\underset{\sim}{c} &= 12\underset{\sim}{k}.5\underset{\sim}{k} \\ &= 60 \end{aligned}$$

So $(\underset{\sim}{a} \times \underset{\sim}{b}).\underset{\sim}{c}$ gives the volume of the rectangular prism.

It is worth considering whether the order of the vectors would affect the result.

$$\begin{aligned} (\underset{\sim}{b} \times \underset{\sim}{c}).\underset{\sim}{a} &= \left(4\underset{\sim}{j} \times 5\underset{\sim}{k}\right).3\underset{\sim}{i} \\ &= 20\underset{\sim}{i}.3\underset{\sim}{i} \\ &= 60 \end{aligned}$$

We will now attempt to solve this problem for three general vectors $\underset{\sim}{a}$, $\underset{\sim}{b}$ and $\underset{\sim}{c}$ (Figure 8.15).

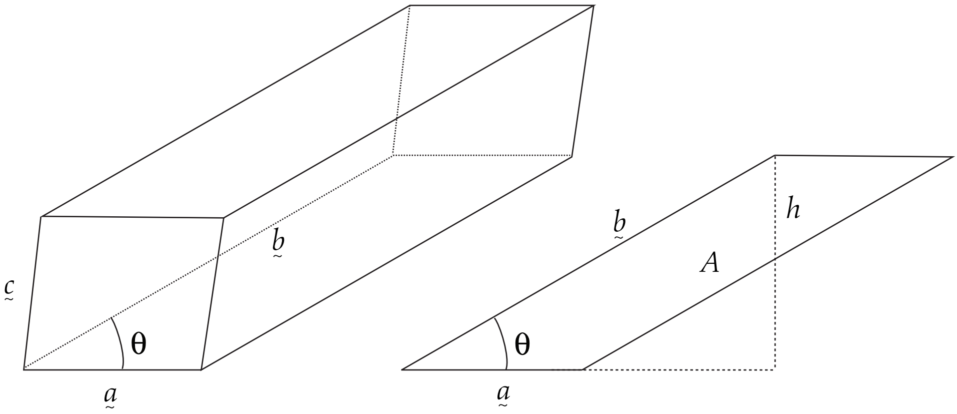


Figure 8.15: Volume of a parallelepiped

$$h = |\underset{\sim}{b}|\sin(\theta) \qquad (1)$$

The area A is $|\underset{\sim}{a}| \times h$.

Using equation (1):

$$A = |\underset{\sim}{a}||\underset{\sim}{b}|\sin(\theta)$$

Interpreting this in terms of the cross product:

$$\underset{\sim}{a} \times \underset{\sim}{b} = |\underset{\sim}{a}||\underset{\sim}{b}|\sin(\theta)\underset{\sim}{d}$$

where $\underset{\sim}{d}$ is a unit vector perpendicular to $\underset{\sim}{a}$ and $\underset{\sim}{b}$.

Let's now consider the scalar triple

$$\begin{aligned}(\underset{\sim}{a} \times \underset{\sim}{b}).\underset{\sim}{c} &= A\underset{\sim}{d}.\underset{\sim}{c} \\ &= A|\underset{\sim}{d}||\underset{\sim}{c}|\cos(\alpha) \\ &= A|\underset{\sim}{c}|\cos(\alpha)\end{aligned}$$

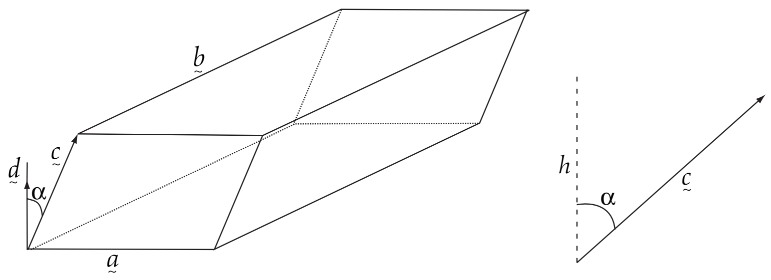


Figure 8.16: Volume of a parallelepiped

As h is the perpendicular height of the box:

$$\therefore (\underset{\sim}{a} \times \underset{\sim}{b}).\underset{\sim}{c} = Ah = \text{volume of box}$$

This means that $(\underset{\sim}{a} \times \underset{\sim}{b}).\underset{\sim}{c}$ gives the volume of a slant box (called a parallelepiped) whose edges are the vectors $\underset{\sim}{a}$, $\underset{\sim}{b}$ and $\underset{\sim}{c}$.

We would get the same result with the following combinations of $\underset{\sim}{a}$, $\underset{\sim}{b}$ and $\underset{\sim}{c}$:

$$(\underset{\sim}{a} \times \underset{\sim}{b}).\underset{\sim}{c} = (\underset{\sim}{b} \times \underset{\sim}{c}).\underset{\sim}{a} = (\underset{\sim}{c} \times \underset{\sim}{a}).\underset{\sim}{b}$$

It is possible to develop a range of questions that directly use both the dot and cross product to solve practical problems. However, to gain a deeper understanding of both forms of multiplication students should explore some of their properties. For this reason Student Activity 8.1 is included.

STUDENT ACTIVITY 8.1

Let $\underset{\sim}{u}$, $\underset{\sim}{v}$ and $\underset{\sim}{w}$ represent three vectors in three-dimensional space and c represent a scalar. Show that:

a $\underset{\sim}{v}.\underset{\sim}{v} = |\underset{\sim}{v}|^2$

b $\underset{\sim}{v}.\underset{\sim}{u} = \underset{\sim}{u}.\underset{\sim}{v}$

c $c(\underset{\sim}{v}.\underset{\sim}{u}) = (c\underset{\sim}{u}).\underset{\sim}{v} = \underset{\sim}{u}.(c\underset{\sim}{v})$

d $\underset{\sim}{v}.(\underset{\sim}{u} + \underset{\sim}{w}) = \underset{\sim}{v}.\underset{\sim}{u} + \underset{\sim}{v}.\underset{\sim}{w}$

e $\underset{\sim}{v}.\underset{\sim}{v} > 0$ if $\underset{\sim}{v} \neq 0$

f $\underset{\sim}{v}.\underset{\sim}{v} = 0$ if and only if $\underset{\sim}{v} = 0$

g $\underset{\sim}{v} \times \underset{\sim}{u} = -\underset{\sim}{u} \times \underset{\sim}{v}$

h $\underset{\sim}{v} \times (\underset{\sim}{u} + \underset{\sim}{w}) = \underset{\sim}{v} \times \underset{\sim}{u} + \underset{\sim}{v} \times \underset{\sim}{w}$

i $(\underset{\sim}{u} + \underset{\sim}{w}) \times \underset{\sim}{v} = \underset{\sim}{u} \times \underset{\sim}{v} + \underset{\sim}{w} \times \underset{\sim}{v}$

j $c(\underset{\sim}{v} \times \underset{\sim}{u}) = (c\underset{\sim}{u}) \times \underset{\sim}{v} = \underset{\sim}{u} \times (c\underset{\sim}{v})$

k $\underset{\sim}{v} \times \underset{\sim}{v} = 0$

l $\underset{\sim}{u}.(\underset{\sim}{u} \times \underset{\sim}{v}) = 0$

m $\underset{\sim}{v}.(\underset{\sim}{u} \times \underset{\sim}{v}) = 0$

n $|\underset{\sim}{u} \times \underset{\sim}{v}|^2 = |\underset{\sim}{u}|^2 |\underset{\sim}{v}|^2 - (\underset{\sim}{u}.\underset{\sim}{v})^2$

o $\underset{\sim}{u} \times (\underset{\sim}{v} \times \underset{\sim}{w}) = (\underset{\sim}{u}.\underset{\sim}{w})\underset{\sim}{v} - (\underset{\sim}{u}.\underset{\sim}{v})\underset{\sim}{w}$

p $(\underset{\sim}{u} \times \underset{\sim}{v}) \times \underset{\sim}{w} = (\underset{\sim}{u}.\underset{\sim}{w})\underset{\sim}{v} - (\underset{\sim}{v}.\underset{\sim}{w})\underset{\sim}{u}$

q $|\underset{\sim}{u} \times \underset{\sim}{v}| = |\underset{\sim}{u}||\underset{\sim}{v}|\sin\theta$

SUMMARY

- There are two methods of multiplying (finding the product of) vectors:
  - the *scalar* or *dot* product
  - the *vector* or *cross* product
- The *scalar* (or *dot*) product is defined as $\underset{\sim}{a}.\underset{\sim}{b} = |\underset{\sim}{a}||\underset{\sim}{b}|\cos(\theta)$
  - The *scalar* product in the unit vector system is given by:
    $\underset{\sim}{a} = x_1\underset{\sim}{i} + y_1\underset{\sim}{j} + z_1\underset{\sim}{k}$
    $\underset{\sim}{b} = x_2\underset{\sim}{i} + y_2\underset{\sim}{j} + z_2\underset{\sim}{k}$
  - $\underset{\sim}{a}.\underset{\sim}{b} = x_1x_2 + y_1y_2 + z_1z_2$
  - If the *scalar* product is equal to zero and neither of the vectors has a magnitude of zero, we know the vectors are perpendicular.
  - The *scalar* product models the work done by a system if one of the vectors represents the force applied to an object and the other represents the object's displacement.
  - The *scalar* product can be used to find parallel and perpendicular components for a vector.
    The parallel vector is $\underset{\sim}{r}_1 = (\hat{\underset{\sim}{P}}.\underset{\sim}{r})\hat{\underset{\sim}{P}}$.
    The perpendicular vector is $\underset{\sim}{r}_2 = (\hat{\underset{\sim}{P}}.\underset{\sim}{r})\hat{\underset{\sim}{P}} - \underset{\sim}{r}$

SUMMARY (Cont.)

- The *cross* product of two vectors is defined as $\underset{\sim}{a} \times \underset{\sim}{b} = |\underset{\sim}{a}||\underset{\sim}{b}|\sin(\theta)\hat{\underset{\sim}{n}}$, where $\hat{\underset{\sim}{n}}$ is a unit vector perpendicular to the plane containing $\underset{\sim}{a}$ and $\underset{\sim}{b}$, and oriented using the right-hand rule.
- The cross product in the unit vector system is given by

$$\underset{\sim}{a} = x_1\underset{\sim}{i} + y_1\underset{\sim}{j} + z_1\underset{\sim}{k}$$

$$\underset{\sim}{b} = x_2\underset{\sim}{i} + y_2\underset{\sim}{j} + z_2\underset{\sim}{k}$$

$$\underset{\sim}{a} \times \underset{\sim}{b} = (y_1z_2 - y_2z_1)\underset{\sim}{i} + (z_1x_2 - x_1z_2)\underset{\sim}{j} + (x_1y_2 - x_2y_1)\underset{\sim}{k}$$

- The cross product has a number of useful applications which include:
  - measuring the turning effect of a force (torque)
  - finding the volume of a parallelepiped.

Websites

http://www.phy.syr.edu/courses/java-suite/crosspro.html
http://www.ee.surrey.ac.uk/Teaching/Courses/EFT/dynamics/html/crossproduct.html

Both of these sites contain an interactive tutorial that describes the cross product and allows students to conduct calculations.

http://www.netcomuk.co.uk/~jenolive/homevec.html

This site contains a number of tutorials and applications for vectors, in particular, applications for both the vector and cross product.

CHAPTER 9

IT'S A CIRCUS

The applications of vectors can be extended if we consider methods that allow vectors to represent curves in space. This can be achieved by the use of parametric functions of time for components of a vector. This representation of vectors allows us to describe the paths of objects as they move in three-dimensional space. It is worth revisiting the context of navigation using parametric functions to describe the paths of ships as they travel across the ocean. We can then extend this context to consider the paths of projectiles, which can be easier for students to visualise.

PLOTTING A PATH

Our previous exploration of the use of vectors to describe the journey of ships across the ocean was limited because the discussion was limited to vectors that described either the *position* of a ship or its *velocity* at a specific time.

When we consider the location of a ship, we know that its position and thus its position vector will vary with time. We can describe this changing position using the parametric equation:

$$\underset{\sim}{s}(t) = x(t)\underset{\sim}{i} + y(t)\underset{\sim}{j}$$

where t is a real number greater than or equal to zero.

We can visualise this path by plotting a series of points for specific values of time (t) (Figure 9.1).

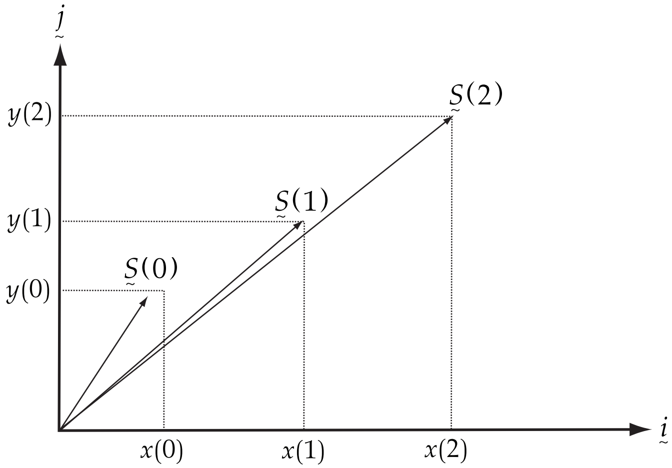


Figure 9.1: A particle moving in a plane

The position of the ship $(\underset{\sim}{s})$ is a vector function of time, t. We know that when we measure time it is a real scalar value. So we can define the position of the ship in three dimensions as a vector function. It is defined by:

$$\underset{\sim}{s}(t) = x(t)\underset{\sim}{i} + y(t)\underset{\sim}{j} + z(t)\underset{\sim}{k}, \ t \geq 0$$

where $x(t)$, $y(t)$ and $z(t)$ are ordinary real functions of time.

A fascinating aspect of these functions is the ability to plot points on a three-dimensional space for individual moments. So, while we are plotting a point(x, y, z), each point is given as a separate function of time. We call the functions the $x = x(t)$, $y = y(t)$and $z = z(t)$ *parametric* equations.

Parametric equations create some interesting possibilities when you consider the paths they describe. We should consider two options:

- finding the cartesian equation of the path and using this equation to find the path of a ship as it moves through a plane
- using modern technology to find the path from the parametric form directly. This process is convenient when using algebra to find a cartesian equation proves to be difficult.

FINDING CARTESIAN EQUATIONS FROM PARAMETRIC EQUATIONS

To understand how to approach finding the cartesian equation of the path of a particle for a given expression for $\underset{\sim}{s}(t)$ we should consider an example.

Consider the vector $\underset{\sim}{r}(t)$, where

$$\underset{\sim}{r}(t) = (t+1)\underset{\sim}{i} + \left(t^2 + t\right)\underset{\sim}{j},\ t \geq 0$$

For $\underset{\sim}{r}(t)$:

$$x = t + 1$$
$$y = t^2 + t$$

Before we find the cartesian equation we should consider the possible values for x and y given $t \geq 0$.

If $\qquad x = t + 1 \qquad (1)$

When $\qquad t \geq 0, x \geq 1$

and for $\qquad y = t^2 + t \qquad (2)$

when $\qquad t \geq 0, y \geq 0$

We can rewrite equation (1) to give $t = x - 1$

Substituting this equation into equation (2) gives:

$$\begin{aligned} y &= (x-1)^2 + (x-1) \\ &= x^2 - 2x + 1 + -1 \\ &= x^2 - x \\ &= x(x-1), x \geq 1 \end{aligned}$$

The path is the part of the parabola $y = x^2 - x$ for which $x \geq 1$ (Figure 9.2).

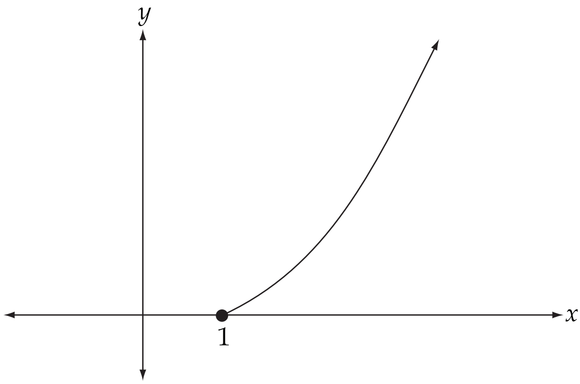


Figure 9.2: The parabola $y = x^2 - x$

Let us now consider the path of a particle with equation

$$\underset{\sim}{s}(t) = \sin(2t)\underset{\sim}{i} + \cos(t)\underset{\sim}{j},\ t \geq 0$$

It would be difficult to find the cartesian form of these parametric equations, but with the use of an appropriate graphics package it is possible to sketch the path of the particle (see Figure 9.3).

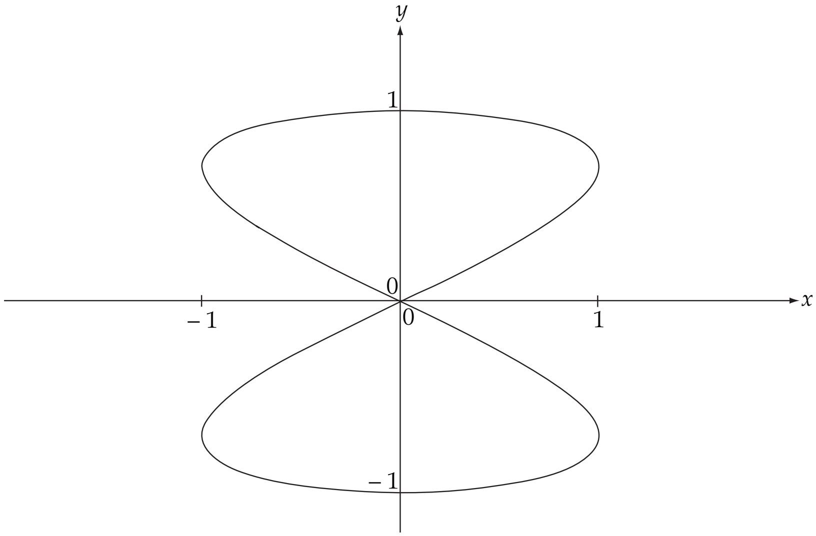


Figure 9.3: Graph of $\underset{\sim}{s}(t) = \sin(2t)\underset{\sim}{i} + \cos(t)\underset{\sim}{j},\ t \geq 0$

VECTOR CALCULUS

Consider again the general equation of the position vector

$$\underset{\sim}{s}(t) = x(t)\underset{\sim}{i} + y(t)\underset{\sim}{j} + z(t)\underset{\sim}{k}$$

We can differentiate $\underset{\sim}{s}(t)$ by differentiating each of the components of the parametric equations separately.

$$\underset{\sim}{s}'(t) = x'(t)\underset{\sim}{i} + y'(t)\underset{\sim}{j} + z'(t)\underset{\sim}{k}$$

The derivative of the position vector gives the velocity vector of a particle with the path $\underset{\sim}{s}(t)$.

The derivative of the velocity vector gives the acceleration vector of a particle with the path $\underset{\sim}{s}(t)$

$$\underset{\sim}{s}''(t) = x''(t)\underset{\sim}{i} + y''(t)\underset{\sim}{j} + z''(t)\underset{\sim}{k}$$

Using vector calculus offers great power when we want to analyse and predict the paths of particles as they travel through space.

SUPPORTING CIRCUS PERFORMERS

Great circus acts require great courage and great skills. At the circus, young and old watch in awe as performers risk their lives for our entertainment. Consider the human cannonball. A person is shot across an open space to land in a net. For the performer it is extremely important to correctly locate the net, if they are to survive the trick unharmed.

When watching the human cannonball, we anticipate the danger. What happens if he or she is shot out of the cannon too quickly or too slowly. He or she may not reach the net or, worse, overshoot it. Our human cannonball knows the risks and so is wise enough to use a mathematician to carefully prepare the act.

MATHEMATICS AND THE HUMAN CANNONBALL

When a cannonball is fired, it starts with an initial velocity, and once it leaves the muzzle the only force that acts on it is gravity.

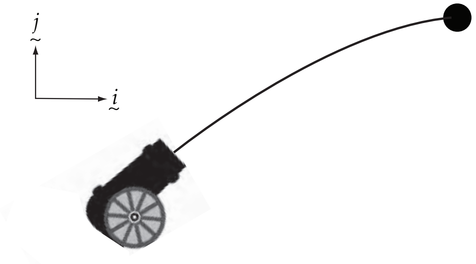


Figure 9.4: The path of the human cannonball

Using vector notation, the force on the cannonball as it leaves the muzzle would be $\underset{\sim}{F} = -mg\underset{\sim}{j}$ where m is the mass of the canonball and g is the acceleration due to gravity.

As the ball leaves the muzzle, its acceleration would be described by the equation $\underset{\sim}{a} = -g\underset{\sim}{j}$.

We know that:

$$\underset{\sim}{a} = \frac{d\underset{\sim}{v}}{dt}$$

$$\frac{d\underset{\sim}{v}}{dt} = -g\underset{\sim}{j}$$

$$\therefore \underset{\sim}{v} = -gt\underset{\sim}{j} + \underset{\sim}{c}$$

To find *c*, we need to know the initial velocity of the cannonball.

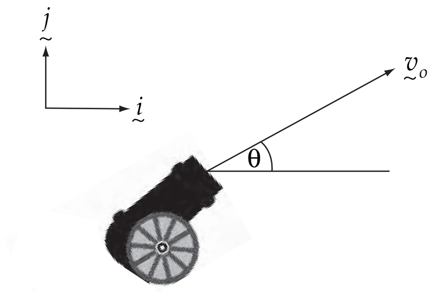


Figure 9.5: Initial trajectory of human cannonball

The initial velocity would be $\underset{\sim}{v}_o = v_o \cos(\theta)\underset{\sim}{i} + v_o \sin(\theta)\underset{\sim}{j}$.

$$\text{At } t = 0,\ \underset{\sim}{v} = \underset{\sim}{c}$$

$$\therefore \underset{\sim}{v} = v_o \cos(\theta)\underset{\sim}{i} + \left(v_o \sin(\theta) - gt\right)\underset{\sim}{j}$$

We can track the position of the path of the ball using $\frac{d\underset{\sim}{s}}{dt} = \underset{\sim}{v}$, where $\underset{\sim}{s}$ is the position of the ball at any time t.

$$\frac{d\underset{\sim}{s}}{dt} = v_o \cos(\theta)\underset{\sim}{i} + \left(v_o \sin(\theta) - gt\right)\underset{\sim}{j}$$

$$\underset{\sim}{s} = v_o t \cos(\theta)\underset{\sim}{i} + \left(v_o \sin(\theta) - \tfrac{1}{2}gt^2\right)$$

We can visualise the path of the cannonball by converting the parametric equations into their cartesian forms.

$$x = v_o t \cos(\theta)$$

$$y = v_o t \sin(\theta) - \tfrac{1}{2}gt^2$$

$$t = \frac{x}{v_o \cos(\theta)}$$

Solving these equations simultaneously:

$$y = \sin(\theta) \times \frac{x}{\cos(\theta)} - \tfrac{1}{2} g \left(\frac{x}{v_o \cos(\theta)} \right)^2$$

$$= x \tan(\theta) - \tfrac{1}{2} g \frac{x^2}{v_o^2 \cos^2(\theta)}$$

$$= x \left(\tan(\theta) - \frac{\tfrac{1}{2} g}{v_o^2 \cos^2(\theta)} x \right)$$

When $y = 0$, $x = 0$ and $x = \dfrac{v_o^2 \sin(2\theta)}{g}$

These equations give us the path of a ball with an initial velocity, v_o and an initial position $(x, y) = (0, 0)$.

Graphically the path would look like that in Figure 9.6.

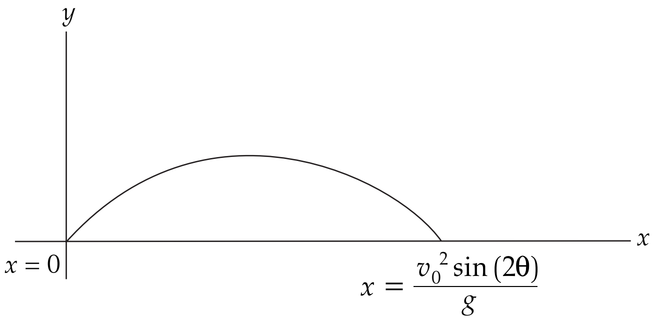


Figure 9.6: Path of projectile

From this graph, it can be seen that the endpoint of the ball's journey is given by:

$$x = \frac{v_o^2 \sin(2\theta)}{g}$$

The distance travelled by the ball is determined by the initial velocity and the angle of the shot. Now for any given velocity, the angle (θ) of the shot will determine the total distance travelled. We know that the maximum value for $\sin(2\theta) = 1$ and that this occurs when $2\theta = 90°$. Thus the maximum distance for a shot will occur when $\theta = 45°$.

The human cannonball now has four equations which can be used to plot the path:

$$\underset{\sim}{a} = -g\underset{\sim}{j}$$
$$\underset{\sim}{v} = v_o\cos(\theta)\underset{\sim}{i} + (v_o\sin(\theta) - gt)\underset{\sim}{j}$$
$$\underset{\sim}{s} = v_o t\cos(\theta)\underset{\sim}{i} + \left(v_o t\sin(\theta) - \tfrac{1}{2}gt^2\right)\underset{\sim}{j}$$
$$y = x\tan(\theta) - \tfrac{1}{2}y\frac{x^2}{{v_o}^2\cos^2(\theta)}$$

These equations can be used to make a number of important decisions that will make his stunt look more exciting but allow the human cannonball to manage the risks involved.

The human cannonball essentially has control of three elements of the stunt:

1 the speed at which he or she exits the cannon
2 the angle at which the cannonball is fired
3 the placement of the net

An aspect worth exploring is the relationship between the initial velocity (v_i) and the angle of the shot (θ).

When the human cannonball is considered as a spectacle, it needs to be realised that the crowd will gasp with excitement if the human cannonball reaches a great height and is in the air for a long time.

Any student of mathematics can find possibilities for the human cannonball, using a graphics calculator or a computer algebra system.

Using parametric equations for the position of the human cannonball:

$$x = v_o t\cos(\theta)$$
$$y = v_o t\sin(\theta) - \tfrac{1}{2}gt^2$$

Using these equation and a variety of values for v_o and θ, appropriate technology can be used to investigate the paths the human cannonball could use to arrive at the net. Explore the options, attempting to balance the excitement of the act and the performer's safety. It is also helpful to take this a step further by considering the location of the net and the speed at which the performer hits the net.

BAREBACK HORSE RIDING

Another circus performance that lends itself to analysis using vector calculus is bareback horse riding. Bareback riders perform a variety of tricks standing the backs of horses as they trot around the circus ring.

During the act a group of clowns arrive with a toy cannon. They set it up in the middle of the circus ring and then proceed to fire volleyballs at the riders. Of course, the riders catch the balls and throw them at the clowns, knocking them off their platform.

We can use mathematics to describe the interaction between rider and volleyball. We will start with the rider. As the horse trots around the ring we will represent the rider as a point (Figure 9.7).

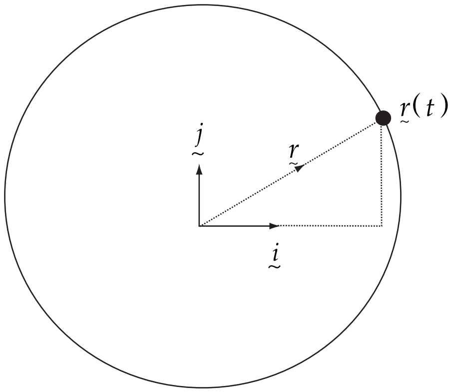


Figure 9.7: Rider is represented as point

The vector equation that represents the path of the bareback rider is $\underset{\sim}{r}(t) = r\cos(nt)\underset{\sim}{i} + r\sin(nt)\underset{\sim}{j},\ t \geq 0$.

The cartesian equation of the path will be gained by solving the parametric equations, $x = r\cos(nt)$ and $y = r\sin(nt)$.

$$\begin{aligned} x^2 + y^2 &= r^2(\cos(nt))^2 + r^2(\sin(nt))^2 \\ &= r^2\left(\cos^2(nt) + \sin^2(nt)\right) \\ \Rightarrow x^2 + y^2 &= r^2 \end{aligned}$$

The velocity of the rider is found by differentiating the position vector $\underset{\sim}{r}(t)$.

$$\begin{aligned} \underset{\sim}{r}(t) &= r\cos(nt)\underset{\sim}{i} + r\sin(nt)\underset{\sim}{j} \\ \underset{\sim}{r}'(t) &= -nr\sin(nt)\underset{\sim}{i} + rn\cos(nt)\underset{\sim}{j} \end{aligned}$$

The rider acceleration would be $\underset{\sim}{r}''(t) = -n^2 r\cos(nt)\underset{\sim}{i} - n^2 r\sin(nt)\underset{\sim}{j}$

Two things worth briefly mentioning:

- $\underset{\sim}{r}'(t).\underset{\sim}{r}''(t) = 0$, which suggests that velocity and acceleration are perpendicular to each other.
- $\underset{\sim}{r}''(t) = -n^2\left(\underset{\sim}{r}(t)\right)$, which some may recognise as the basic formula for simple harmonic motion.

We should return to the clowns and their volleyballs. To solve this problem we need to add the third dimension.

For a rider to safely catch the ball it should hit the rider's chest, which is about 2 metres above the ground.

$$r(t) = r\cos(nt)\underset{\sim}{i} + r\sin(nt)\underset{\sim}{j} + 2\underset{\sim}{k}$$

We can use our previous work with the human cannonball with some re-orientation of the axes because we will use $\underset{\sim}{k}$ rather than $\underset{\sim}{j}$ as the vertical plane.

The equation for the volleyball is

$$s = v_o t\cos(\theta)\underset{\sim}{i} + \left(v_o t\sin(\theta) - \tfrac{1}{2}gt\right)^2\underset{\sim}{k}$$

The question that needs to be considered is: when do the clowns fire the cannon so that the ball arrives in time for it to be caught?

To make the solution of this problem easier, we will orientate the axes so the path of the volleyball is parallel to the y-axis.

This means we need to find the time it takes the ball to reach the point $\underset{\sim}{s} = r\underset{\sim}{i} + 0\underset{\sim}{j} + 2\underset{\sim}{k}$ from the point $\underset{\sim}{s} = 0\underset{\sim}{i} + 0\underset{\sim}{j} + 0\underset{\sim}{k}$.

To solve this problem we need to consider a pair of simultaneous equations:

$$v_o t\cos(\theta) = r \text{ and } v_o t\sin(\theta) - \tfrac{1}{2}yt^2 = 2$$

As you can imagine, there are any number of solutions, depending on our choice of the initial velocity, the angle of the shot and the radius of the circus ring.

Once again it is possible to explore values for r, θ and v_o using technology.

Once values for r, θ and v_o have been set we can find the time of flight of the ball. We can use this value to find how far the rider will travel on the back of the horse.

There are many options that can be explored using the scenario of clowns shooting at horse riders. All we need to realise is that the equations we have used do not change much, only the initial conditions change.

Once we have set the initial condition we can restate our equations to solve problems that involve question about position and time.

STUDENT ACTIVITY 9.1

The *Ferris Wheel* is a challenging problem (see Board of Studies, 1999). It is included here because it gives an indication of the types of tasks that can be developed for student as projects that use vector parametric equations to describe events that may be familiar to them.

Throughout this problem, assume air resistance is negligible, and take the acceleration due to gravity to have magnitude g m/s^2, where $g = 9.8$.

Peter goes to the Royal Show, and takes a ride on the ferris wheel. While on this ride he can be represented as a point.

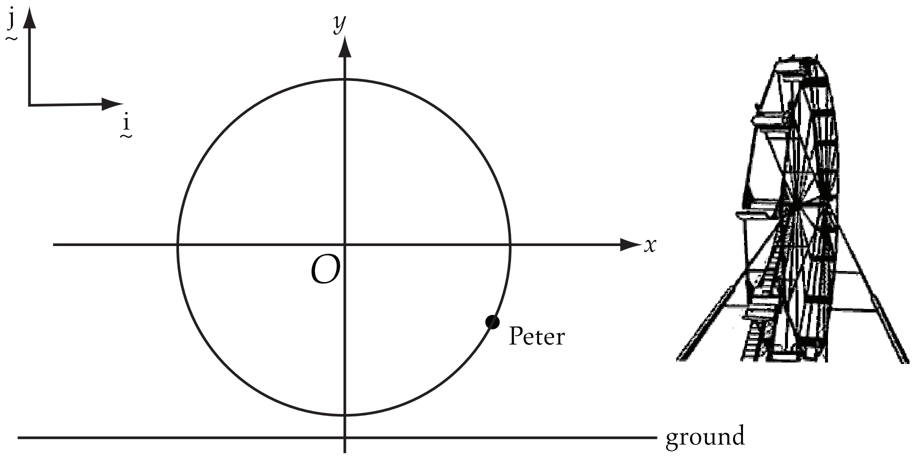


Figure 9.8: Ferris wheel problem

The Ferris wheel has a radius of 15 metres, and rotates anticlockwise in a vertical plane, at a constant rate, around a horizontal axis located at its centre. Its axis is located 18 metres above the ground. Peter's position at any time t, in seconds, is given by:

$$\underset{\sim}{P}(t) = 15\cos\left(\frac{\pi t}{60}\right)\underset{\sim}{i} + 15\sin\left(\frac{\pi t}{60}\right)\underset{\sim}{j}$$

where Peter's initial position is 15 metres to the right of the vertical axis and 18 metres above the ground.

1 a How long does it take Peter to return to his starting point for the first time?
  b What are the coordinates of Peter's exact position at $t = 45$?
  c Find the cartesian equation of Peter's path.

2 While he is on the ferris wheel, Peter drops a ball at $t = 45$.
  a i What is the horizontal velocity of the ball at the moment it is dropped?
    ii What is the vertical velocity of the ball at the moment it is dropped?

b i Find an expression for the horizontal position of the ball as it falls to the ground, in terms of T_d, where T_d represents the time in seconds from the moment Peter drops the ball.

ii Find an expression for the vertical position of the ball as it falls to the ground, in terms of T_d.

Jonathan, Peter's friend, is standing on the ground. He catches the ball and then throws it back to Peter. When Jonathan throws the ball, he releases it at the point (–16, –16) with a velocity of u m/s and at an initial angle of trajectory θ degrees, $0 < \theta < 90$, measured in an anticlockwise direction from the positive x-direction. Let T_b represent the time in seconds from the moment Jonathan releases the ball.

3 a Find an expression in terms of T_b and the constants u and θ, for:

i the horizontal velocity of the ball

ii the vertical velocity of the ball

b i Find an expression for the x-coordinate of the ball in terms of T_b, and the constants u and θ.

ii Find an expression for the y-coordinate of the ball in terms of T_b and the constants u and θ.

c Show that the cartesian equation of the ball's path is

$$y = -16 + (x+16)\tan(\theta) - \tfrac{1}{2}g\left(\frac{x+16}{u\cos(\theta)}\right)^2$$

Peter catches the ball thrown by Jonathan. To catch the ball, Peter must be at the same position as the ball at the same time.

4 a Just as the ball reaches the top of its path, Peter is at the point (–15, 0) and catches the ball. Establish the values of u and θ correct to two decimal places.

b Given that Peter catches the ball thrown by Jonathan, what were the coordinates of Peter's position, correct to two decimal places, at the moment Jonathan threw the ball?

Jonathan throws another ball to Peter. Peter catches the ball when he arrives at the point (–l5, 0).

5 Given that the ball follows the path described by the cartesian equation

$$y = -16 + (x+16)\tan(\theta) - \tfrac{1}{2}g\left(\frac{x+16}{u\cos\theta}\right)^2$$

and that it must reach the point (–15, 0), express u in terms of θ, where u and θ are as previously defined.

SUMMARY

- Vectors functions can define the position of a point in three-dimensional space. These functions have the form $\underset{\sim}{s}(t) = x(t)\underset{\sim}{i} + y(t)\underset{\sim}{j} + z(t)\underset{\sim}{k},\ t \geq 0$.
- The functions $x = x(t)$, $y = y(t)$ and $z = z(t)$ are called parametric functions.
  - Parametric functions can be converted to relations involving cartesian equations by setting up simultaneous equations and eliminating the variable t from these equations.
  - When converting to relations involving cartesian equations, particular care needs to be given to the domain of the function. Students need to remember that the values that t can hold may restrict the values that x may hold.
- Vector calculus
  - Vector functions can be differentiated and antidifferentiated with respect to t:

$$\underset{\sim}{s}(t) = x(t)\underset{\sim}{i} + y(t)\underset{\sim}{j} + z(t)\underset{\sim}{k},\ t \geq 0$$

$$\underset{\sim}{s}'(t) = x'(t)\underset{\sim}{i} + y'(t)\underset{\sim}{j} + z'(t)\underset{\sim}{k},$$

$$\underset{\sim}{s}''(t) = x''(t)\underset{\sim}{i} + y''(t)\underset{\sim}{j} + z''(t)\underset{\sim}{k},\ t \geq 0$$

  - Vector calculus is used to analyse and predict the paths of particles as they travel through space.
  - Two applications of vector calculus are projectile motion and circular motion.

Websites

http://www.physicsclassroom.com

This site gives further insight into vectors and projectile motion. It includes some animations.

http://jersey.uoregon.edu/newCannon/nc1.html

This site contains a ballistic simulator which students could use to test their calculations.

CHAPTER 10

IT'S NOW POSSIBLE TO KNOW WHERE WE ARE!

The global positioning system (GPS) offers an exciting and interesting application by which students could be introduced to the use of matrices to represent vectors. The GPS creates the opportunity for students to realise that vectors do not need to be limited to three dimensions and thus we can have an n-dimensional space. Applications and the workings of the GPS could be used as a project for students.

In this chapter the application of the GPS as a means of finding your location is explored at its simplest level; however, it may be possible for some students to explore the more complex areas of mathematics suggested in the final paragraphs of the chapter.

THE IMPORTANCE OF KNOWING WHERE YOU ARE

Previously in Chapter 6, we heard the story of Admiral Sir Clowdisley Shovell and his small fleet. The loss of two thousand lives was due to not knowing their precise location on a fog-shrouded ocean. The joint efforts of navigators and navy captains failed to pinpoint their latitude and longitude. This failure meant the course they plotted put them in harm's way.

Can you imagine the awe that Admiral Shovell would have for a GPS—a system that allows anyone with a hand-held device to know where they are, regardless of weather conditions? It would be a dream come true for Admiral Shovell to have a simple device that not only gave him his exact location, but also allowed him to plot a course.

The GPS uses 24 satellites to pinpoint the three-dimensional location of a GPS receiver. The system was developed and deployed by the United States of

America's Department of Defence. Its initial development was designed to give USA defence forces the advantage of quickly knowing their exact location in the 'fog of war'.

Sadly, it took the destruction of a Korean Air 747, with the loss of passengers and crew, to make the GPS available to all. Korean Air Flight 007 was shot down by USSR jets because through a navigational error the 747 had invaded Russian air space. This occurred in 1983, when the Cold War was at its height.

Again we were faced with the loss of life because a captain did not know his exact location. However, this time the world was offered a solution that would allow anyone who could afford to purchase a GPS receiver to know their location with remarkable accuracy.

WORKING OUT YOUR LOCATION

It is worth developing an appreciation of how the satellites are used by a GPS receiver to pinpoint its location. To achieve such an appreciation we need to understand the mechanics and mathematics of global positioning.

The mechanics of global positioning

Each of the 24 satellites continuously transmits two sets of signals. These signals report the exact location of a satellite and the exact time—'GPS' time given by the satellite's atomic clock. The receiver uses this information to establish the distance between the satellite and the receiver.

If we have the exact distance between a satellite and a receiver we can draw a sphere. The centre of the sphere will be located at the satellite. Its radius is the distance between the satellite and the receiver (Figure 10.1).

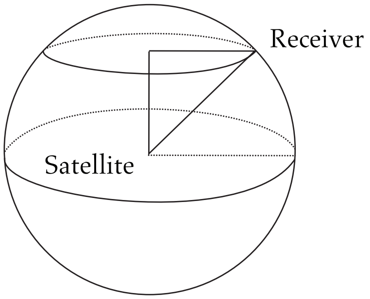


Figure 10.1: A sphere indicating the possible position of a receiver

When the receiver makes contact with a second satellite it now has a second sphere with a known centre and radius. This additional information means that our location will be on the circle of the two intersecting spheres.

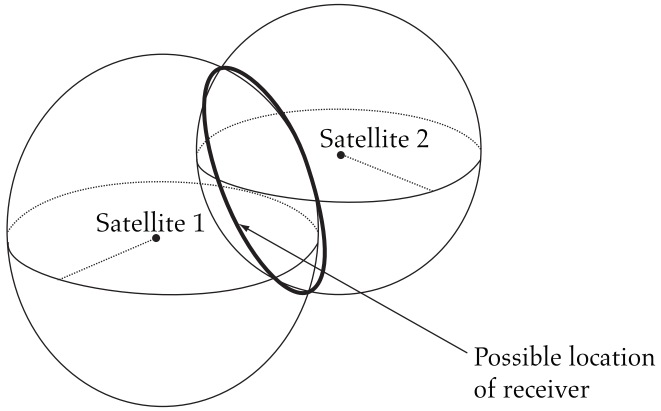


Figure 10.2: Intersecting spheres indicating the possible location for a receiver

To find the exact location of the GPS receiver you need a third satellite. This allows you to find the exact location of the two points at which the three spheres will intersect.

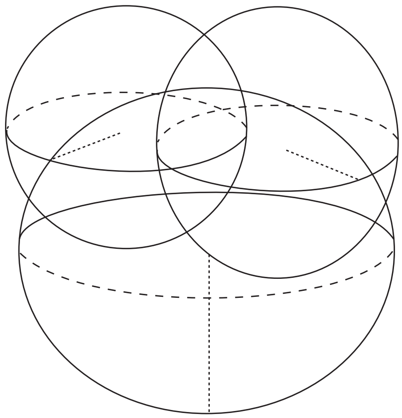


Figure 10.3: Three intersecting spheres

Three intersecting spheres have two points in common.

We know one of the points is not possible, because it is located off the Earth's surface. With this information it is possible to know the exact position of a receiver.

We do need to be careful. The system relies on the ability to establish the exact distance from three satellites based on the length of time it takes for a signal from space to reach a hand-held device. The difficulty of finding this time can be highlighted when we realise that the signals travel at the speed of light. At this speed the signal would travel around the Earth eight times in one second. This places in context the importance of the atomic clocks.

GPS receivers do not contain atomic clocks, so we need a fourth satellite to establish the time. Once the receiver has estimated its position using three satellites, it matches the possible positions with the time signal of a fourth satellite. We know that the (x, y, z) position that has been reasoned can only match the time signal of the fourth satellite at one precise time. So with the addition of a fourth satellite we can find our location.

The fourth satellite creates a mathematical challenge. In a sense we need to use vector with an extra dimension. Previously we have used vectors in three dimension $\underset{\sim}{r} = x\underset{\sim}{i} + y\underset{\sim}{j} + z\underset{\sim}{k}$. We now need to consider a vector that has a fourth dimension, time.

The addition of a fourth dimension can create a complication for vector calculation or it can create an opportunity for us to think about vector operations in a more imaginative way, which can lead to n-dimensional space. To realise the possibility we will turn to matrix representation of vectors.

VECTORS AND VECTOR OPERATIONS USING MATRICES

This will be a quick tour of matrices to represent and operate with vectors. It is not intended to be a comprehensive exploration of matrices. A greater insight and exploration can be gained by reading *Matrices*, in this series.

Any vector $\underset{\sim}{a} = x\underset{\sim}{i} + y\underset{\sim}{j} + z\underset{\sim}{k}$ can be represented as a column matrix:

$$\underset{\sim}{a} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

It is possible to define a vector as a set of real numbers arranged in a column:

$$\underset{\sim}{a} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

In a sense we are now no longer limited to considering vectors as three-dimensional. The dimension of any vector space can be as large as we desire. This allows us the ability to define an n-dimensional space. For this reason it is worth defining a vector space, which is the basic object of the study in the branch of mathematics called linear algebra.

A vector space is a non-empty set of objects, called elements, that satisfy 10 axioms (http://www.cs.caltech.edu). Let V denote a vector space.

Closure axioms:

- Axiom 1—Closure under addition. For every pair of elements x and y in V there corresponds a unique element in V called the sum of x and y, denoted by $x + y$.
- Axiom 2—Closure under multiplication by real numbers. For every x in V and every real number a there corresponds an element in V called the product of a and x, denoted by ax.

The addition, subtraction and scalar multiplication of vectors follows the rules established in Chapter 7.

To add or subtract vectors we simply add or subtract the elements in the same position in the matrices.

$$\text{If } \underset{\sim}{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ . \\ . \\ . \\ a_n \end{bmatrix} \text{ and } \underset{\sim}{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ . \\ . \\ . \\ b_n \end{bmatrix}$$

$$\underset{\sim}{a} + \underset{\sim}{b} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \\ . \\ . \\ . \\ a_n + b_n \end{bmatrix} \text{ and } \underset{\sim}{a} - \underset{\sim}{b} = \begin{bmatrix} a_1 - b_1 \\ a_2 - b_2 \\ a_3 - b_3 \\ . \\ . \\ . \\ a_n - b_n \end{bmatrix}$$

We can also multiply a vector by a scalar k:

$$k\underset{\sim}{a} = \begin{bmatrix} ka_1 \\ ka_2 \\ ka_3 \\ . \\ . \\ . \\ ka_n \end{bmatrix}$$

It is a relatively simple step to realise that it is possible to add any number of vectors of the same size.

$$\underset{\sim}{y} = \underset{\sim}{a} + \underset{\sim}{b} + \underset{\sim}{c} + \ldots \underset{\sim}{z}$$

It is probably worth mentioning the norm of a vector at this point. The norm of a vector is a real number that represents its geometric length. This is easy to imagine in two or three dimensions, because it would be considered to be the length of the vector. It becomes a little more difficult to visualise this property when we move beyond three dimensions. To be a norm it must satisfy three properties:

- The norm of every vectors is positive.
- Scaling a vector scales the norm by the same amount.
- The norm of the sum of two vectors is less than or equal to the sum of the norms of the individual vectors (http://cnx.rice.edu/content/m10768/latest/).

LINEAR COMBINATION OF VECTORS

We can take this a step further by also considering scalar multiplication of vectors. Using addition, subtraction and scalar multiplication we can create new vectors that are linear combination of given vectors.

$$\underset{\sim}{a} = \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix} \text{ and } \underset{\sim}{b} = \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}$$

Consider two linear combination of these vectors:

$$\begin{aligned} \underset{\sim}{c} &= \underset{\sim}{a} + 2\underset{\sim}{b} \\ &= \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix} + \begin{bmatrix} 10 \\ 8 \\ 2 \end{bmatrix} = \begin{bmatrix} 14 \\ 10 \\ 0 \end{bmatrix} \\ \underset{\sim}{d} &= -\underset{\sim}{a} + 3\underset{\sim}{b} \\ &= \begin{bmatrix} -4 \\ -2 \\ +2 \end{bmatrix} + \begin{bmatrix} 15 \\ 12 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 10 \\ 5 \end{bmatrix} \end{aligned}$$

All the possible linear combinations of a set of vectors is called its *linear span*.

In each case above, $\underset{\sim}{c}$ and $\underset{\sim}{d}$ are linear combination of the vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$.

We can generalise this to define linear dependence as a combination of vectors such that

$$a_1 \underset{\sim}{u}_1 + a_2 \underset{\sim}{u}_2 + a_3 \underset{\sim}{u}_3 + \ldots + a_n \underset{\sim}{u}_k = 0$$

where a_n is a non-zero scalar and $\underset{\sim}{u}_k$ is a vector.

If none of the vectors in a set of vectors can be expressed as a linear combination of the remaining vectors, we know that the vectors are linearly independent.

TRANSPOSITION OF VECTORS

Vectors can also be represented as a row matrix rather than a column matrix. A row matrix is the transposition of the column matrix, by simply turning it on its side. It is represented by $\underset{\sim}{a}^T$.

$$\text{If } \underset{\sim}{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ . \\ . \\ . \\ a_n \end{bmatrix} \text{ then } \underset{\sim}{a}^T = \begin{bmatrix} a_1 & a_2 & a_3 & \ldots & a_n \end{bmatrix}$$

We use $\underset{\sim}{a}^T$ to calculate the scalar product for a pair of vectors that are represented as matrices.

If we have two vectors:

$$\underset{\sim}{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ . \\ . \\ . \\ a_n \end{bmatrix} \qquad \underset{\sim}{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ . \\ . \\ . \\ b_n \end{bmatrix}$$

$$\underset{\sim}{a}.\underset{\sim}{b} = a^T \times b = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ . \\ . \\ . \\ b_n \end{bmatrix}$$

$$= a_1 b_1 + a_2 b_2 + a_3 b_3 + \dots + a_m b_m$$

USING MATRIX MULTIPLICATION

We can also use matrix multiplication to define the cross product of two vectors. First let us complete a quick review of finding the determinant of a matrix.

For a square matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

The determinant of $A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21}$

For example:

$$A = \begin{bmatrix} 1 & 6 \\ 3 & 2 \end{bmatrix}$$

$$|A| = \begin{vmatrix} 1 & 6 \\ 3 & 2 \end{vmatrix} = 2 - 18 = -16$$

We can now extend this process to a 3 × 3 matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{aligned} |A| &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{32}a_{23}a_{11} \end{aligned}$$

The determinant of a 3 × 3 matrix offers a simple way of calculating the cross product. It can be defined as the determinant of the matrix.

$$\begin{bmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

$$\begin{aligned} \underset{\sim}{a} \times \underset{\sim}{b} &= \begin{vmatrix} \underset{\sim}{i} & \underset{\sim}{j} & \underset{\sim}{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \\ &= \underset{\sim}{i}(a_2b_3 - a_3b_2) + \underset{\sim}{j}(a_1b_3 - a_3b_1) + \underset{\sim}{k}(a_1b_2 - a_2b_1) \end{aligned}$$

SOLVING SIMULTANEOUS EQUATIONS

A wonderful application of matrices is the solution of simultaneous equations. This is particularly true with the use of appropriate technology.

We know that some students can struggle with solving simultaneous equations with two unknowns. Many of us do not venture into solving simultaneous equations with three, four or more unknowns, because it is time consuming. Worse, if we make a slight error, which can easily occur, the answers we generate will be incorrect.

Using matrices and suitable technology we can become adventurous with the number of equations and unknowns with which we can successfully work. The problem is no longer solving the equations but, rather, understanding and setting up the simultaneous equations.

Let's explore and learn the technique, by starting with a problem with two unknowns.

Consider the following pair of simulataneous equations:

$$3x - 2y = -1 \quad (1)$$

$$6x + 2y = 0 \quad (2)$$

The first step is to construct a coefficient matrix of the variables. Each column of the coefficent matrix represents a variable, while each row represents an equation. This means that the above simultaneous equations can be rewritten as a matrix equation.

$$\begin{bmatrix} 3 & -2 \\ 6 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

If we can find the inverse of the first matrix, then the solution is relatively easy.

If $A = \begin{bmatrix} 3 & -2 \\ 6 & 2 \end{bmatrix}$, the solution to the equation would be $\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}\begin{bmatrix} -1 \\ 0 \end{bmatrix}$

Before we solve this equation, there is an interesting aside which may appear obvious after we have started it. We know that if we plot a pair of simultaneous equations on a set of cartesian coordinates the solution will be the intersection point of the lines.

A special case occurs when the lines are parallel. In this case the lines will not intersect and so there will be no solution to the simultaneous equations.

For example:

$$\begin{aligned} 2y + 4x &= 12 \\ y + 2x &= 13 \end{aligned}$$

If we set this up as a coefficient matrix:

$$\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 12 \\ 13 \end{bmatrix}$$

However, we would not expect that this matrix to have an inverse such that:

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}\begin{bmatrix} 12 \\ 13 \end{bmatrix}$$

Our previous work on finding the determinant gives us a quick way of discovering if a matrix has an inverse.

In this case $\begin{vmatrix} 1 & 4 \\ 2 & 2 \end{vmatrix} = 0.$

If the determinant of a matrix is zero then the matrix does not have an inverse. A matrix that has a determinant of zero is called *singular.*

This knowledge is extremely useful where we are attempting to find if the vectors in a system are *linearly dependent* or *independent*. In a similar way to our previous simultaneous equations, vector that are linearly dependent are collinear. This means we can use the determinant of a coefficient matrix that describes a system of vectors to discover if they are linearly dependent.

Consider the following simple example:

$$\underset{\sim}{r}_1 = 3\underset{\sim}{i} + 4\underset{\sim}{j}$$
$$\underset{\sim}{r}_2 = 8\underset{\sim}{i} + 2\underset{\sim}{j}$$

This system can be represented as:

$$\begin{bmatrix} 3 & 4 \\ 8 & 2 \end{bmatrix}\begin{bmatrix} \underset{\sim}{i} \\ \underset{\sim}{j} \end{bmatrix}$$
$$\begin{vmatrix} 3 & 4 \\ 8 & 2 \end{vmatrix} = 6 - 32 = -26$$
$$\begin{vmatrix} 3 & 4 \\ 8 & 2 \end{vmatrix} \neq 0$$

therefore these vector are linearly independent.

However, if the determinant of a matrix that represents a system of vectors is singular then the vectors are linearly dependent.

Referring back to our simultaneous equations:

$$\begin{bmatrix} 3 & -2 \\ 6 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}\begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Using appropriate technology we will find that:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{-1}{9} \\ \frac{1}{3} \end{bmatrix}$$

We can use this technique to solve simultaneous equation with 3, 4 or more unknowns.

$$2x - 3y + 7z = -1$$
$$1.5x + 0.5y + 3z = 0$$
$$2x - 10y = 4$$

becomes

$$\begin{bmatrix} 2 & -3 & 7 \\ 1.5 & 0.5 & 3 \\ 2 & 10 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}\begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 0.6 \\ 0.4 \end{bmatrix}$$

MATHEMATICS AND THE LOCATION OF EARTHQUAKES

Before we consider the mathematics of global positioning, it is worth exploring how earthquakes are located. The mathematics is similar to that for global positioning, but for earthquakes we need only consider two dimensions.

When an earthquake occurs, seismic stations record the shock waves that are produced. There are two types of shock waves. The primary waves travel relatively quickly through the earth and are the first to arrive at a seismic station. A secondary wave is also generated by the earthquake and will arrive some time after the primary wave.

Seismologists know the speed at which both the primary and secondary waves travel through the earth. They use this to establish the distance of the earthquake from the seismic station.

During a recent earthquake, three seismic stations recorded the arrival of both the primary and secondary shock waves. Analysis of these waves placed the epicentre of the earthquake 1000 km from station A, 2500 km from station B and 1751 km from station C.

The location of station A is (1000, 1000), station B is (1700, 1280) and station C is (700, 1480) from a specific marker. The location of the earthquake is (x, y).

The vectors from the epicentre to each seismic station would be:

$$\underset{\sim}{A}:\ (x - 1000, y - 1000)$$
$$\underset{\sim}{B}:\ (x - 1700, y - 1280)$$
$$\underset{\sim}{C}:\ (x - 700, y - 1480)$$

We know the length of each of these vectors so we can create three corresponding equations using this knowledge.

$$(x-1000)^2+(y-1000)^2=(1000)^2$$
$$(x-1700)^2+(y-1280)^2=(2500)^2$$
$$(x-700)^2+(y-1480)^2=(1751)^2$$

These equations expand to:

$$x^2-2000x+1000^2+y^2-2000y+1000^2=1000^2$$
$$x^2-3400x+1700^2+y^2-2650y+1280^2=2500^2$$
$$x^2-1400x+700^2+y^2-2960y+1480^2=1751^2$$

These equations can be simplified:

$$x^2-2000x+y^2-2000y=-(1000)^2$$
$$x^2-3400x+y^2-2560y=1\,721\,600$$
$$x^2-1400x+y^2-2960y=385\,601$$

We should now subtract two pairs of equations:

$$\text{first} - \text{second: } 1400x+560y=-2\,721\,600$$
$$\text{second} - \text{third: } -2660x+400y=1\,335\,999$$

These two equations can be set up as a coefficient matrix:

$$\begin{bmatrix} 1400 & 450 \\ -2660 & 400 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} -2\,721\,600 \\ 1\,335\,999 \end{bmatrix}$$
$$\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} -1093.333 \\ -2126.6675 \end{bmatrix}$$

We can use a similar technique to locate a position using global positioning. The mathematics we would use is as follows.

We can start by assuming that the GPS receiver contains an atomic clock that is synchronised with the atomic clocks on the satellites. Using this assumption each satellite would give us an equation of the form:

$$(x-\delta x_i)^2+(y-\delta y_i)^2+(z-\delta z_i)^2=c^2\,t_i$$

δx_i, δy_i, δz_i and t_i differ for each satellite and represent the coordinates and the time taken for the message to reach the GPS receiver respectively. The speed of light is represented by the constant c.

We can now use this information to write a set of simultaneous equations, whose solution will give us the exact position of the receiver.

Oh, that it were that simple. It is more difficult. There are two major reasons why the process becomes more complicated. First, the satellites are not static and, second, we cannot easily establish the time delay, because the the speed of light is so high the time it takes a message to arrive at a receiver will be extremely small.

Hence, the mathematics becomes quite complex. While it is beyond the constraints of this book, it is worth our time to look at the algorithm used. The actual mathematics has been described on the website http://everything2.com and is as follows.

For each satellite we have an equation of the form:

$$(x_1 - sx_n)^2 + (x_2 - sy_n)^2 + (x_3 - sz_n)^2 + c^2 (sd_n - x_4)^2 = 0$$

sx, sy, sz and sd differ for each satellite and they represent the three coordinates and the measured time delay respectively; c is the speed of light. We need to solve the position (x_1, x_2, x_3), as well as the *error* in the receiver clock x_4.

The algorithm used for this purpose is as follows. It is taken directly from http://everything2.com.

1 Arrange the satellite equations into a system $S_{n,j}(x)$ with $(n, j) = 1 \ldots 4$. If we have five equations we can construct a coefficient matrix by subtracting one of the equations from the rest.
2 Make an initial guess for positions x_1, x_2, x_3 and the time correction factor x_4. Choosing a position at the origin and a correction so as not to cause any negative delays is appropriate.
3 Calculate the Jacobian matrix $J_{n,j}$ for S_n with $(n, j) = 1 \ldots 4$ more equations result in an over-determined system. There are 24 satellites in the GPS systems.
4 Solve the $n \times n$ linear system $Jy = -S$ for vector y, where we have substituted into $S(x)$ our initial guess for x.
5 Update the variables: $x_{\text{new}} = x_{\text{old}} + y$.
6 Check for convergence: $x_{\text{new}}^2 - x_{\text{old}}^2 <$ tolerance, otherwise return to step 2.

This problem becomes more complex when you realise that the the arrival of the signal may be affected, if only slightly, by the location of the receiver. Receivers located in canyons, for example, may be affected by delays because the signal cannot travel directly to the receiver.

It's now possible to know where we are!

SUMMARY

- Global positioning systems (GPS) create an opportunity to explore vectors that are not limited to three dimensions.
  - GPS uses 24 satellites to pin-point the location of a GPS receiver.
  - GPS required a receiver to receive a signal from 4 satellites.
  - The time taken for a signal to reach the receiver from three satellites is used to identify three intersecting spheres and thus the two possible locations of the receiver. The fourth satellite is used to establish the GPS time for the receiver. One of the locations of the receiver is eliminated because it would not be on the Earth's surface.
- We use matrices to represent vectors because it eases the required calculations.
- The introduction of matrices as a representation of vectors allows students to be introduced to the concept of vector space and n-dimensional space.
- The four operations of vectors established in earlier chapters are consistent with matrix operations. We can use the properties of finding the determinant of a vector to define linear dependence, and perform vector and scalar products.
- Matrices offer an efficient method for solving simultaneous equations. It is this method of solving simultaneous equations that allows us a ready method for locating earthquakes or a GPS receiver. This method is made up of three steps.
- The simultaneous equations $ax + by = c$ and $dx + ey = f$ are represented by:

$$\begin{bmatrix} a & b \\ d & e \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c \\ f \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ d & e \end{bmatrix}^{-1}\begin{bmatrix} c \\ f \end{bmatrix}$$

Websites

http://www.everything2.com/
This website contains information about the mathematics of global positioning as well as a series of links which offer greater detail.

http://www.cs.caltech.edu/~westside/quantum-glossary.htm
This site contains a number useful definitions, one which—the definition of a vector space—was used in this chapter.

http://cnx.rice.edu/content/m10768/latest/
This is another site that contains useful definitions of vector concepts.

CHAPTER 11

PONS ASINORUM—THE ASSES' BRIDGE

It seems appropriate to end this book (metaphorically) where we began, in ancient Greece, and in particular with geometry. This chapter is structured to illustrate a range of proofs of geometric propositions using vector methods. Teachers can select a variety of these as well as others to illustrate the benefits of including vector methods in a student's repertoire of proof techniques. Such work can be supplemented by the use of dynamic geometry software which can be used to investigate the constructions behind the proofs.

Similar to the application of vectors, geometry can be used to calculate lengths, areas and volumes. The Babylonians and Egyptians, mostly through trial and error, developed a large functional knowledge base that was used for construction, navigation and surveying. In fact the Babylonians knew a great deal about navigation and astronomy, and their knowledge was based on a good understanding of geometry. They were the first to speculate that the earth was a sphere.

The geometry of the Egyptians and Babylonians found a home in ancient Greece. However, the Greeks refused to accept without question geometric statements discovered through trial and error, as such a method of discovery may mean that some of the computation is wrong. Thales of Milete insisted that any geometric statement be established by using deductive reasoning. This notion of proof made Greek mathematics a distinctive and powerful analytical tool.

Euclid, a member of the Platonic school, gave himself the task of collecting the work of Thales, Pythagoras and his disciples, and Hippocrates into a central text: *Euclid's Elements*. This simple act allowed him to become the most widely read author in history. The approach he laid down has dominated the teaching of geometry in the West for more than two millennia. It would be worth encouraging students to research the influence of Euclid on the teaching of mathematics. Some references are included that could act as a starting point.

His influence is felt beyond geometry. The method he used to prove geometric statements is the process that is central to pure mathematics. It is pure because it is achieved through our imagination only (pure thought). It is not sullied by the need for physical experimentation to verify results.

This formal approach to mathematics is often called the axiomatic method. The axioms are explicit statements of assumptions, together with any definitions and other statements such as theorems. Lemmas and corollaries must be shown to follow logically from the axioms and definitions.

Geometry was used by some ancient Greeks to divide people into two groups: those who could understand at least the fifth proposition of Book 1 of *Euclid's Elements* and those who couldn't. This theorem states that the angles at the base of an isosceles triangle are equal. Anyone who could move beyond this proposition was not an ass. He had crossed the bridge of asses. I believe we could move forward to proposition 47 of Book 1 of *Euclid's Elements* for our pons asinorum.

Are you wondering and hoping that you could cross our asses' bridge? All that is required is to prove the geometric statement: 'In right-angled triangles the square on the side opposite the right angle equals the sum of the squares on the sides containing the right angle.' This proposition is often called Pythagoras' theorem. We won't use Euclid's proof because it relies on earlier propositions. However, it is possible to use vector algebra to assist in establishing proposition 47. Before we travel this path we should consider vector proofs.

It is worth noting that there are distinct differences in the various approaches to 'proof' in geometry. The proofs used by Euclid use synthetic geometry because they use theorems and synthetic observations to draw conclusions. These differ from analytical geometry, which uses algebra to perform geometric operations and thus solve problems. Vectors offer a third alternative for some geometric proofs. Student sometimes mix methods from synthetic geometry, coordinate geometry and vectors when presenting a proof, which can lead to confusion.

VECTOR PROOFS

We can prove many geometric statements using position vectors. In this case, each point on a two-dimensional plane can be associated with a vector. This approach covers some of the key ideas in mathematics and geometry. An example is the division of a line segment in a given ratio. This can be achieved with a great deal of elegance using vectors.

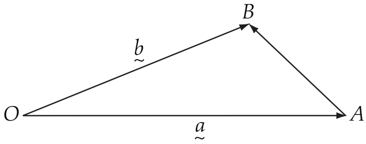
We can achieve this by fixing an origin, O, defining position vectors to a point A as the vector $\overrightarrow{OA}$. Previously we used $\underset{\sim}{a}$ to represent this vector. We will use this as a principle for vector proofs.

If we want to continue in a similar style to that required of Euclid's student we should start with a simple lemma.

Lemma 1

Given two position vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ then $\overrightarrow{AB} = \underset{\sim}{b} - \underset{\sim}{a}$.

Proof:



$$\begin{aligned}\overrightarrow{AB} &= \overrightarrow{AO} + \overrightarrow{OB}\\ &= -\overrightarrow{OA} + \overrightarrow{OB}\\ &= -\underset{\sim}{a} + \underset{\sim}{b}\\ &= \underset{\sim}{b} - \underset{\sim}{a}\end{aligned}$$

Using a simple definition for the division of a vector into a given ratio, it is possible to create sufficient mathematical tools to prove some of the properties of triangles.

Definition:

The point Q on a vector $\overrightarrow{AB}$ divides it into the ratio $u{:}v$ where $u + v \neq 0$ if $u\overrightarrow{AQ} = v\overrightarrow{QB}$.

Using this definition we can create a formula to divide any vector $\overrightarrow{AB}$ in a given ratio.

Lemma 2

If Q divides a vector $\overrightarrow{AB}$ into the ratio $u{:}v$ and $u + v \neq 0$, then the position vector of Q is

$$\underset{\sim}{P} = \frac{u}{u+v}\underset{\sim}{a} + \frac{v}{u+v}\underset{\sim}{b}$$

Proof:

We know by definition

$$u\overrightarrow{AB} = v\vec{B}$$

This suggests

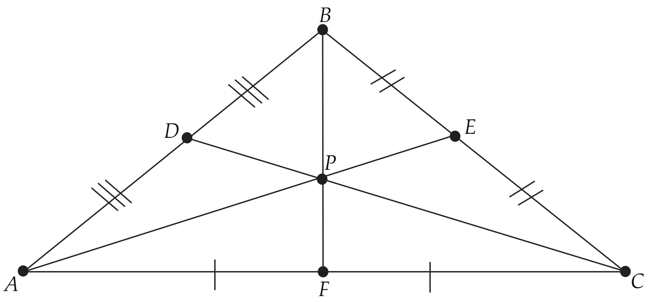
$$a(\underset{\sim}{P} - \underset{\sim}{a}) = v(\underset{\sim}{b} - \underset{\sim}{P})$$
$$u\underset{\sim}{P} - u\underset{\sim}{a} = v\underset{\sim}{b} - v\underset{\sim}{P}$$
$$u\underset{\sim}{P} + v\underset{\sim}{P} = u\underset{\sim}{a} + v\underset{\sim}{b}$$
$$\underset{\sim}{P}(u + b) = u\underset{\sim}{a} + v\underset{\sim}{b}$$
$$\underset{\sim}{P} = \frac{u}{u+v}\underset{\sim}{a} + \frac{v}{u+v}\underset{\sim}{b}$$

We can now use this result to prove that the medians of a triangle are concurrent. In other words, the medians of a triangle all meet at the same point. This point is known as the *centroid* of a triangle.

Theorem 1

For a triangle with vertices *A*, *B* and *C*, the medians of *ABC* are concurrent. This point (the centroid), *P*, divides the medians in the ratio 2:1.

Proof:



Let *D* be the midpoint of *AB*
E be the midpoint of *BC*
F be the midpoint of *AC*

By lemma 2

$$\underset{\sim}{d} = \tfrac{1}{2}\underset{\sim}{a} + \tfrac{1}{2}\underset{\sim}{b}$$

Let P divide the line in the ratio 2:1.
Again using lemma 2

$$\begin{aligned}\underset{\sim}{P} &= \tfrac{1}{3}\underset{\sim}{c} + \tfrac{2}{3}\underset{\sim}{d} \\ &= \tfrac{1}{3}\underset{\sim}{c} + \tfrac{2}{3}\left(\tfrac{1}{2}\underset{\sim}{a} + \tfrac{1}{2}\underset{\sim}{b}\right) \\ &= \tfrac{1}{3}\underset{\sim}{a} + \tfrac{1}{3}\underset{\sim}{b} + \tfrac{1}{3}\underset{\sim}{c}\end{aligned}$$

This formula is symmetric for A, B and C. This means we have the same point if we started with E or F.

Therefore P lies on all three medians.

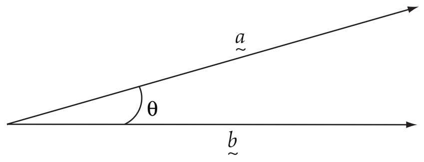
USING THE SCALAR PRODUCT

Although we have previously defined the scalar product, we will restate it.

Definition:

$$\underset{\sim}{a}.\underset{\sim}{b} = |\underset{\sim}{a}||\underset{\sim}{b}|\cos(\theta)$$

where $|\underset{\sim}{a}|$ and $|\underset{\sim}{b}|$ is the magnitude of the vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ respectively. The angle θ is the angle subtended by the vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$.

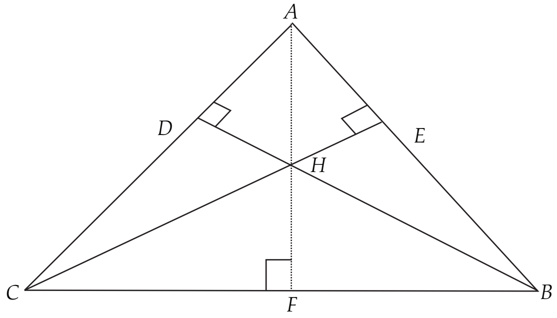


The scalar product can be used to prove several theorems.

Theorem 2: The orthocentre of a triangle

An altitude of a triangle is the line drawn from a vertex and perpendicular to the opposite side. The three altitudes are concurrent. This point is called the *orthocentre* of a triangle.

Proof:



Assume H is the point of intersection $\overrightarrow{FA}$ and $\overrightarrow{DB}$.

We know that:

$\overrightarrow{AH}$ is orthogonal to $\overrightarrow{BC}$
$\overrightarrow{BH}$ is orthogonal to $\overrightarrow{AC}$

The dot product of orthogonal vectors is zero.

We have two equations:

$$(\underset{\sim}{h} - \underset{\sim}{a})(\underset{\sim}{c} - \underset{\sim}{b}) = 0 \quad (1)$$

$$(\underset{\sim}{h} - \underset{\sim}{b})(\underset{\sim}{c} - \underset{\sim}{a}) = 0 \quad (2)$$

(2) – (1):

$$\begin{aligned}
&(\underset{\sim}{h} - \underset{\sim}{a})(\underset{\sim}{c} - \underset{\sim}{b}) - (\underset{\sim}{h} - \underset{\sim}{b})(\underset{\sim}{c} - \underset{\sim}{a}) = 0 \\
&= \underset{\sim}{h}.\underset{\sim}{c} - \underset{\sim}{h}.\underset{\sim}{b} - \underset{\sim}{a}.\underset{\sim}{c} + \underset{\sim}{a}.\underset{\sim}{b} - [\underset{\sim}{h}.\underset{\sim}{c} - \underset{\sim}{h}.\underset{\sim}{a} - \underset{\sim}{b}.\underset{\sim}{c} + \underset{\sim}{b}.a] \\
&= \underset{\sim}{h}.\underset{\sim}{c} - \underset{\sim}{h}.\underset{\sim}{b} - \underset{\sim}{a}.\underset{\sim}{c} + \underset{\sim}{a}.\underset{\sim}{b} - \underset{\sim}{h}.\underset{\sim}{c} + \underset{\sim}{h}.\underset{\sim}{a} + \underset{\sim}{b}.\underset{\sim}{c} - \underset{\sim}{b}.a \\
&= \underset{\sim}{h}.\underset{\sim}{b} + \underset{\sim}{b}.\underset{\sim}{c} - \underset{\sim}{h}.\underset{\sim}{c} - \underset{\sim}{a}.\underset{\sim}{c} \\
&= \underset{\sim}{h}(\underset{\sim}{a} - \underset{\sim}{b}) - \underset{\sim}{c}(\underset{\sim}{a} - \underset{\sim}{b}) \\
&= (\underset{\sim}{h} - \underset{\sim}{c})(\underset{\sim}{a} - \underset{\sim}{b})
\end{aligned}$$

So $\overrightarrow{CH}.\overrightarrow{BA} = 0$

Therefore $\overrightarrow{CH}$ is perpendicular to $\overrightarrow{AB}$.

This means that the third altitude $\overrightarrow{EC}$ passes through the point of intersection of $\overrightarrow{AF}$ and $\overrightarrow{BD}$.

A second proof allows an understanding of the *circumcentre* of a triangle.

The circumcentre of triangle is the point where the vertices A, B and C are equidistant from that point.

We will define H such that

$$\underset{\sim}{h} = \underset{\sim}{a} + \underset{\sim}{b} + \underset{\sim}{c}$$

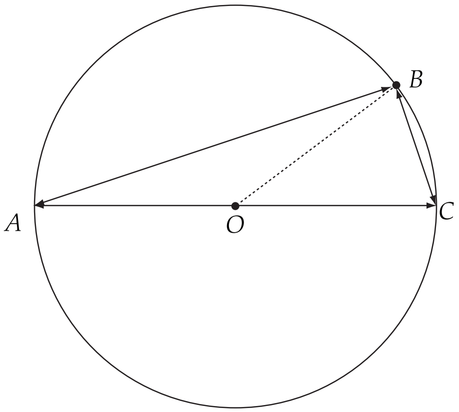
We can show that H lies on each of the altitudes of the triangle ABC.

$$\begin{aligned}
\overrightarrow{AH}.\overrightarrow{BC} &= (-\underset{\sim}{a} + \underset{\sim}{h})(-\underset{\sim}{b} + \underset{\sim}{c}) \\
&= (-\underset{\sim}{a} + \underset{\sim}{a} + \underset{\sim}{b} + \underset{\sim}{c})(\underset{\sim}{c} - \underset{\sim}{b}) \\
&= (\underset{\sim}{b} + \underset{\sim}{c})(\underset{\sim}{c} - \underset{\sim}{b}) \\
&= \underset{\sim}{c}^2 - \underset{\sim}{b}^2 \\
\overrightarrow{OC} &= \overrightarrow{OB} \\
\therefore c^2 - b^2 &= 0
\end{aligned}$$

This process can be repeated for the remaining altitudes.

Theorem 3: Thales' theorem

This notion leads neatly to Thales' theorem: If A, B and C are points on a circle where the line AC is a diameter of the circle then the angle ABC is a right angle.



An interesting approach to prove Thales' theorem is to consider the converse. That is, the hypotenuse of a right triangle is a diameter of its circumcentre.

Proof:

$$\begin{aligned}
\overrightarrow{BA} &= \underset{\sim}{a} - \underset{\sim}{b} \\
\overrightarrow{BC} &= \underset{\sim}{c} - \underset{\sim}{b} \\
\overrightarrow{BA}.\overrightarrow{BC} &= (\underset{\sim}{a} - \underset{\sim}{b})(\underset{\sim}{b} - \underset{\sim}{c})
\end{aligned}$$

From the diagram it can be seen that $\underset{\sim}{b} - \underset{\sim}{c} = \underset{\sim}{a} + \underset{\sim}{b}$

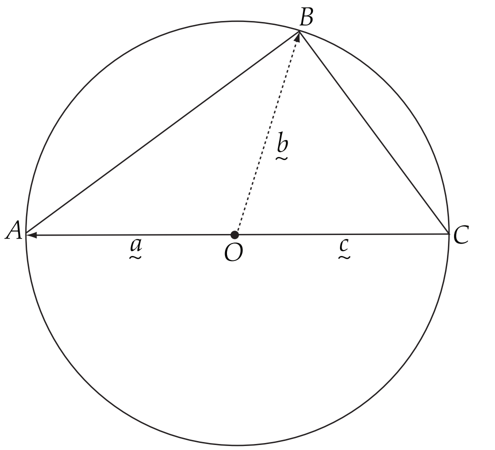
$$\begin{aligned} \text{Hence } \overrightarrow{BA}.\overrightarrow{BC} &= (\underset{\sim}{a} - \underset{\sim}{b})(\underset{\sim}{a} + \underset{\sim}{b}) \\ &= \underset{\sim}{a}.\underset{\sim}{a} - \underset{\sim}{b}.\underset{\sim}{b} \\ &= |\underset{\sim}{a}| - |\underset{\sim}{b}| = 0 \end{aligned}$$

Therefore $|\underset{\sim}{a}| = |\underset{\sim}{b}|$.

A and *B* are also the same distance from the point *O*, the circle's centre. This means that *O* is the circumcentre of the triangle. This will only occur if the angle *ABC* is a right angle.

Theorem 4: Pythagoras' theorem

It is time to cross our asses' bridge and establish proposition 47 or Pythagoras' theorem: In a right-angled triangle, the square on the side opposite the right angle equals the sum of the squares on the sides containing the right angle.



Using lemma 1:

$$\overrightarrow{BA} = \underset{\sim}{a} - \underset{\sim}{b}$$
$$\overrightarrow{BC} = \underset{\sim}{c} - \underset{\sim}{b}$$
$$\overrightarrow{AC} = \underset{\sim}{c} - \underset{\sim}{a}$$

Using Thales' theorem:

$$\overrightarrow{BC} = \underset{\sim}{a} + \underset{\sim}{b}$$

$$\overrightarrow{BA}.\overrightarrow{BA} = \left|\overrightarrow{BA}\right|^2 = (\underset{\sim}{a} - \underset{\sim}{b})(\underset{\sim}{a} - \underset{\sim}{b})$$

$$= \underset{\sim}{a}^2 - 2\underset{\sim}{a}.\underset{\sim}{b} + \underset{\sim}{b}^2$$

$$\overrightarrow{BC}.\overrightarrow{BC} = \left|\overrightarrow{BC}\right|^2 = (\underset{\sim}{a} + \underset{\sim}{b})(\underset{\sim}{a} + \underset{\sim}{b})$$

$$= \underset{\sim}{a}^2 + 2\underset{\sim}{a}.\underset{\sim}{b} + \underset{\sim}{b}^2$$

$$\left|\overrightarrow{BC}\right|^2 + \left|\overrightarrow{BA}\right|^2 = a^2 + 2ab + b^2 + a^2 - 2ab + b^2$$

$$= 2a^2 + 2b^2$$

$$\overrightarrow{AC}.\overrightarrow{AC} = \left|\overrightarrow{AC}\right|^2 = (\underset{\sim}{c} - \underset{\sim}{a})(\underset{\sim}{c} - \underset{\sim}{a})$$

$$= c^2 - 2\underset{\sim}{a}.\underset{\sim}{c} + a^2$$

$\underset{\sim}{a} + \underset{\sim}{c}$ are collinear

$$= \underset{\sim}{c}^2 + 2|a\|c| + \underset{\sim}{a}^2$$

We know that $|\underset{\sim}{a}\|\underset{\sim}{b}| = |\underset{\sim}{c}|$

This means that $\underset{\sim}{c}^2 + 2|\underset{\sim}{a}\|\underset{\sim}{c}| + \underset{\sim}{a}^2$ can be rewritten as $2\underset{\sim}{b}^2 + 2\underset{\sim}{a}^2$

$$\therefore \left|\overrightarrow{BA}\right|^2 + \left|\overrightarrow{BC}\right|^2 = \left|\overrightarrow{AC}\right|^2$$

It may be a good feeling to cross the asses' bridge. However, this crossing is a little easier than for those who cross without using vectors. Vectors provide powerful tools that can simplify the proof of theorems which can otherwise appear daunting. However, care needs to be taken to ensure that the original proposition to be proven, or an equivalent, has not been implicitly assumed in any suggested proof. For Pythagoras' theorem this is particularly important since the 'length' of a vector is often defined in terms of an equivalent form of the theorem. The first website listed below contains useful discussion in relation to the use of vectors in this context (see also Stillwell 2005).

Vector proof is a wonderful introduction to mathematical proof which reflects the approach encouraged by Euclid, but without some of the complexity. It is also worth noting that once we have established a theorem, it can be used to aid future proofs.

Students should be encouraged to attempt to construct proofs of geometric propositions, and this can be done in conjunction with work on making conjectures and testing their validity. The following could be used to give students the opportunity to try out vector proofs. They should be careful and should, as far as possible, use the appropriate conventions of mathematics.

STUDENT ACTIVITY 11.1

Use vectors to prove the following geometric results:

- a The diagonals of a rhombus are perpendicular.
- b The diagonals of a parallelogram bisect each other.
- c When the midpoints of the sides of a rectangle are joined, a rhombus is formed.
- d The midpoint of the hypotenuse of a right triangle is equidistant from the three vertices.
- e The median of the non-equal side of an isosceles triangle is perpendicular to that side.
- f If the diagonals of a parallelogram are of equal length, then the parallelogram is a rectangle.
- g Show that for any triangle $a^2 = b^2 + c^2 - 2bc\cos(\theta)$.

SUMMARY

- The ancient Greeks applied deductive reasoning to the geometry of the Egyptians and Babylonians and thus established the notion of proof to mathematics. This made mathematics a powerful analytical tool.
- Euclid collected the works of Thales, Pythagoras and Hippocrates into central text: *Euclid's Elements*. This text became one of the most widely read texts in history.
- Euclid's method formed a key part of the foundations of pure mathematics in geometry.
- The formal approach began with the explicit statement of definitions and fundamental assumptions, or *axioms*. *Lemmas* (results to assist in the proof) and *corollaries* (subsequent results from the proof) must be shown to follow logically from the axioms and definitions.
- To enable vector proofs of geometric theorems to follow a similar format, we use position vectors to locate specific points in space. With the addition of two simple lemmas we have a powerful analytical tool to prove geometric theorems.
  - Lemma 1: Given two position vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ then $\overrightarrow{AB} = \underset{\sim}{b} - \underset{\sim}{a}$.

SUMMARY (Cont.)

- Lemma 2: If the point Q divides a vector $\overrightarrow{AB}$ into the ratio $u{:}v$ where $u + v \neq 0$, then the corresponding position vector is $\underset{\sim}{Q} = \frac{u}{u+v}\underset{\sim}{a} + \frac{v}{u+v}\underset{\sim}{b}$

- The scalar product can be used as a definition and hence to establish the orthocentre of a triangle, and Thale's theorem. The latter theorem then can be used to prove Pythagoras' theorem.
- Students should be made aware that there are several approaches to a mathematical proof (for example, synthetic geometry, analytical geometry and vector proof). They should not mix these methods when presenting a proof.

Reference

Stillwell, J 2005, *The four pillars of geometry*, Springer-Verlag, New York.

Websites

http://www.cut-the-knot.com/pythagoras/PTcom3.shtml
    This site has a discussion of the use of vectors in establishing the Pythagorean theorem.

http://geometryalgorithms.com/history.htm
    This site gives a brief outline of the history of geometry, paying particular attention to some of the major geometers responsible for its development.

http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Euclid.html
    This site provides a good biography of Euclid and offers many useful links to other sites related to his work, in particular his elements.

CHAPTER 12

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 is 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 used 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 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 types 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 12.1 provides a mapping in terms of curriculum connections between the chapters of this book, the four types of courses identified above, and the courses currently 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 syllabuses, in a particular state or territory.

Table 12.1: Curriculum connections for senior secondary final year mathematics courses in Australia

State or territory	Type of course	Relevant chapters
Victoria	4: Specialist Mathematics	all
New South Wales	4: Mathematics Extension 2	1, 2, 3, 4 and 5
Queensland	4: Mathematics C	all
South Australia & Northern Territory	4: Specialist Mathematics	1, 2, 3, 4 and 5
Western Australia	4: Calculus	all
Tasmania	4: Mathematics Specialised	1, 2, 3, 4 and 5

Table 12.2 provides a mapping in terms of curriculum connections between the chapters of this book, the four types of courses identified above, and some of the courses currently offered in various English-speaking jurisdictions from around the world. Again, as this book is a teacher resource, these connections relate to the usefulness of material from the chapters in terms of mathematical background of relevance, rather than a direct mapping to curriculum content, or syllabuses, in a particular jurisdiction.

Table 12.2: 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	4: Advanced Placement Calculus BC	none
International Baccalaureate Organisation (IBO)	4: Mathematics HL	all
UK	4: Advanced level	all

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 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 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/

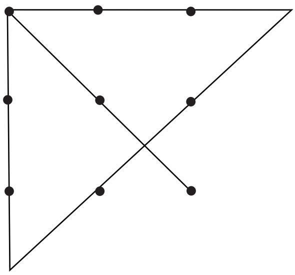
OECD Program for International Student Assessment (PISA)
http://www.pisa.oecd.org

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

CHAPTER 13

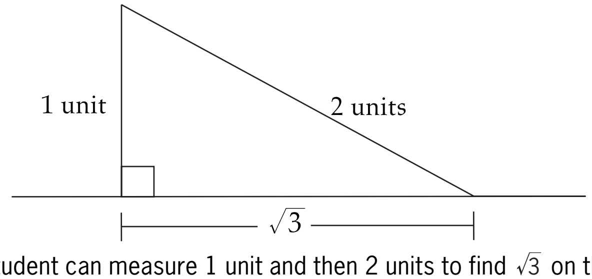
SOLUTION NOTES TO STUDENT ACTIVITIES

Student activity 2.1



Student activity 2.2

Students can attack these in a variety of ways. It is possible to use previous results. It is important, though, to use right angles and the ability to measure units to move forward.



For example, a student can measure 1 unit and then 2 units to find $\sqrt{3}$ on the number line.

Student activity 2.3

a $2n^2 = 4n^2 = 2(2n^2)$

b $(2n+1)^2 = 4n^2 + 4n + 1 = 2(2n^2 + 2) + 1$

c $2m \times 2n = 4mn = 2(2mn)$

d $(2n+1)(2m+1) = 4mn + 2n + 2m + 1 = 2(2mn + n + m) + 1$

e $(2n+1)(2m) = 2m(2n+1)$

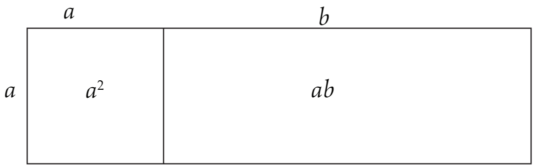
f an odd number: $2n + 1 + 2m = 2(n+m) + 1$

g an even number: $2n + 1 + 2m + 1 = 2n + 2m + 2 = 2(n + m + 1)$

h an even number: $2m + 2n + 2(m+n)$

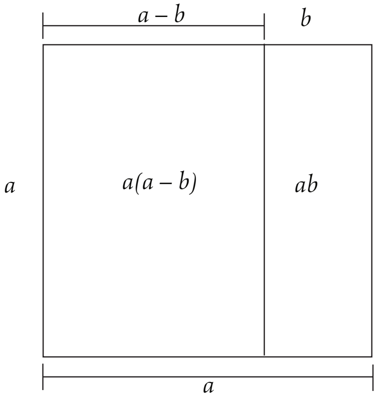
Student activity 2.4

a



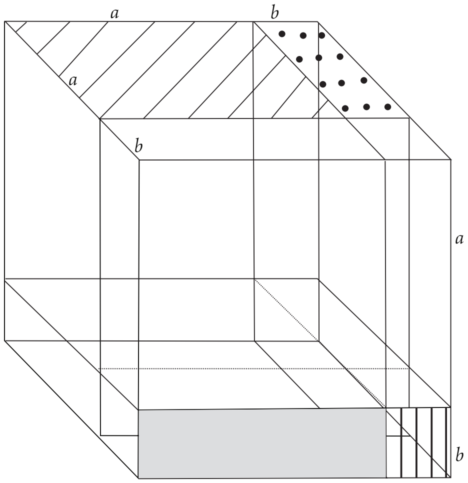
$a(a+b) = a^2 + ab$

b



$a^2 - ab = a(a-b)$

c



Drawing in three dimensions is more difficult to see, but it helps if you add some colour.

Block a^3

Block b^3

Block $3a^2b$

Block $3ab^2$

Total volume $a^3 + 3a^2b + 3ab^2 + b^3$

Student activity 3.1

a $x^3 - 15x^2 + 81x - 175$

$$m = \frac{-63}{2},\ a = 20$$

Depressed form:

$$x^3 - \frac{63x}{2} = 20$$

$$x = 4 + 3i,\ 4 - 3i,\ 7$$

b $x^3 + 8x^2 + 25x + 26$

$$m = -7,\ a = \frac{74}{27}$$

Depressed form:

$$x^3 - 7x = \frac{74}{27}$$

$$x = -3 \pm 2i,\ -2$$

c $2x^3 - 21x^2 + 68x - 29$

$$m = \frac{-169}{4},\ a = \frac{-75}{2}$$

Depressed form:

$$x^3 - \frac{169x}{4} = -\frac{75}{2}$$

$$x = 5 \pm 2i,\ \tfrac{1}{2}$$

d $2x^3 - 25x^2 + 102x - 130$

$$m = -\frac{217}{4},\ a = -\frac{305}{54}$$

Depressed form:

$$x^3 - \frac{217x}{4} = -\frac{305}{54}$$

$$x = 5 \pm i,\ \tfrac{5}{2}$$

Student activity 3.2

a $z^2 + 4z + 6 = 0 \quad z = -2 \pm \sqrt{2}\,i$

b $z^2 + 2z + 8 = 0 \quad z = -1 \pm \sqrt{7}\,i$

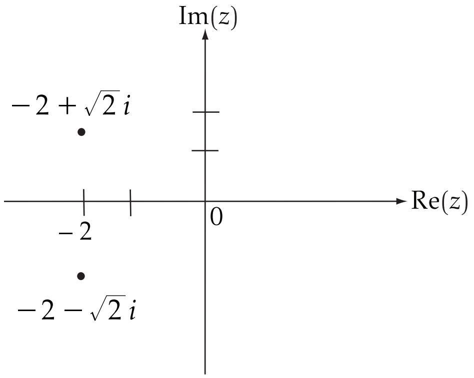
c $z^2 - 2z + 8 = 0 \quad z = 1 \pm \sqrt{7}\,i$

d $z^2 + 4z + 10 = 0 \quad z = 2 \pm \sqrt{6}\,i$

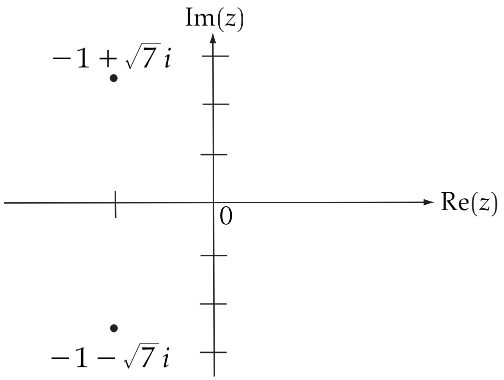
Solutions to e–k can be checked using a computer algebra system.

Student activity 3.3

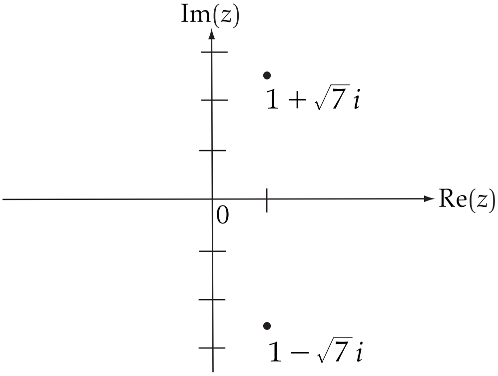
a



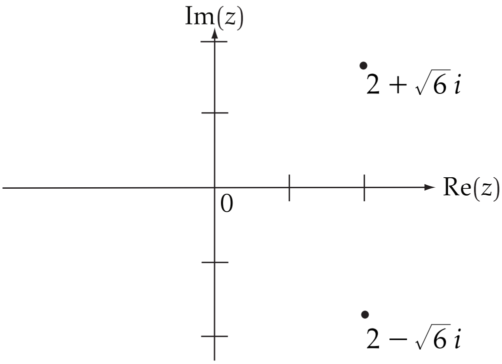
b



c



d



Students should comment that each solution is symmetrical about the *x*-axis.

Student activity 4.1

Results

a i $5 - i$

a ii $5 + \sqrt{3} + 4i$

a iii $-4 - i$

b i For ai

$$|3 - 2i| = \sqrt{13}$$
$$|5 - i| = \sqrt{26}$$

aii and aiii would use a similar approach.

b ii This activity allows students to practise adding complex numbers. One conclusion they should draw is something similar to:

For two complex numbers z_1 and z_2

$$|z_1| + |z_2| \geq |z_1 + z_2|$$

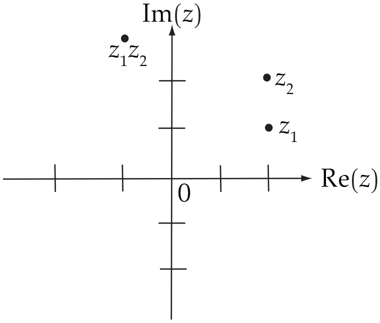
Student activity 4.2

a $z_1 z_2 = -\sqrt{3} + 1 + (\sqrt{3} + 1)$

$z_1 z_3 = \sqrt{3} - 1 + (\sqrt{3} + 1)$

$z_2 z_3 = 4i$

b i



b ii $|z_1| = \sqrt{2}$, $|z_2| = 2$

$|z_1 z_2| = 2\sqrt{2}$

c and d would follow similar processes

e Students may identify the relationship between the modules and argument. That is:
For $z_1 z_2$

$|z_1 z_2| = |z_1||z_2|$

$\arg(z_1 z_2) = \arg z_1 + \arg z_2$

However, they may not express it using these terms.

Student activity 4.3

a $z_1 = \sqrt{2}\,\text{cis}\left(\frac{\pi}{4}\right)$

$z_2 = 2\text{cis}\left(\frac{\pi}{3}\right)$

$z_3 = 2\text{cis}\left(\frac{\pi}{6}\right)$

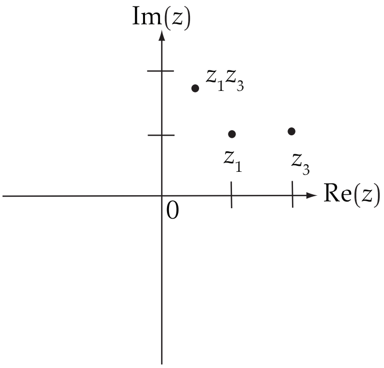
b $z_1 z_2 = 2\sqrt{2}\,\text{cis}\left(\frac{7\pi}{12}\right)$

$z_1 z_3 = 4\text{cis}\frac{5\pi}{12}$

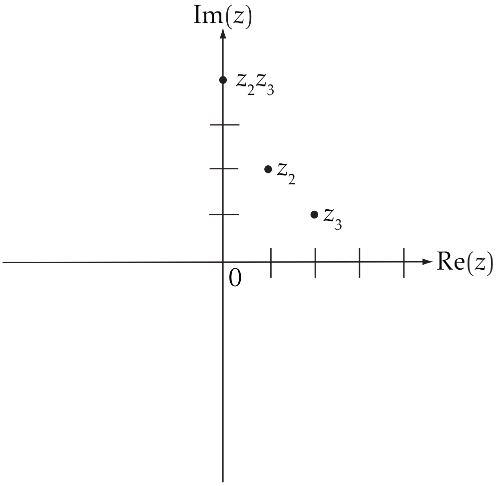
$z_2 z_3 = 4\text{cis}\frac{\pi}{2}$

c See 4.2 bi.

d



e



Student activity 4.4

The most straightforward approach is to express $z = x + iy$. Separate both sides and perform the operation in the LHS or RHS individually.

For example:

$$
\begin{aligned}
z &= x + iy \\
\bar{z} &= x - iy \\
z\bar{z} &= (x + iy)(x - iy) \\
&= x^2 - ixy + ixy - i^2 y^2 \\
&= x^2 + y^2 \\
\text{LHS} &= \text{RHS}
\end{aligned}
$$

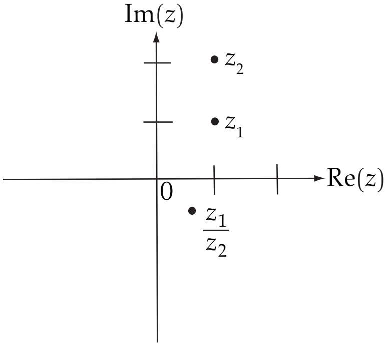
Student activity 4.5

a $\frac{z_1}{z_2} = \frac{\sqrt{3}+1}{4} + \frac{1-\sqrt{3}}{4}i$

$\frac{z_1}{z_3} = \frac{\sqrt{3}+1}{4} + \frac{\sqrt{3}-1}{4}i$

$\frac{z_2}{z_3} = \frac{\sqrt{3}}{2} + \frac{i}{2}$

b i



b ii $|z_1| = \sqrt{2}$ $|z_2| = 2$

$\left|\frac{z_1}{z_2}\right| = \frac{\sqrt{2}}{2}$

b iii Angle for z_1 is $\frac{\pi}{4}$

Angle for z_2 is $\frac{\pi}{3}$

Angle for $\frac{z_1}{z_2}$ is $-\frac{\pi}{12}$

c and d will follow a similar pattern.

Student activity 4.6

a $z_1 = \sqrt{2}\,\text{cis}\left(\frac{\pi}{4}\right)$

$z_2 = 2\text{cis}\left(\frac{\pi}{3}\right)$

$z_3 = 2\text{cis}\left(\frac{\pi}{6}\right)$

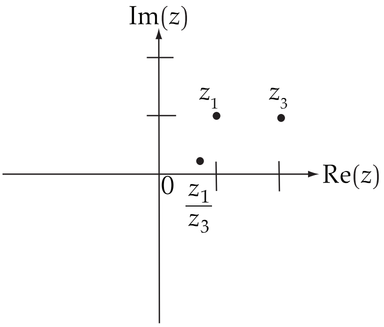
b $\frac{z_1}{z_2} = \frac{\sqrt{2}}{2}\text{cis}\left(\frac{-\pi}{12}\right)$

$\frac{z_1}{z_3} = \frac{\sqrt{2}}{2}\text{cis}\left(\frac{\pi}{12}\right)$

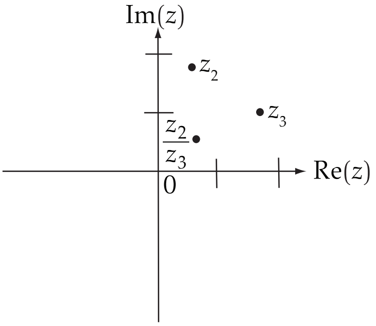
$\frac{z_2}{z_3} = \text{cis}\left(\frac{\pi}{6}\right)$

c See Student Activity 4.5 bi.

d

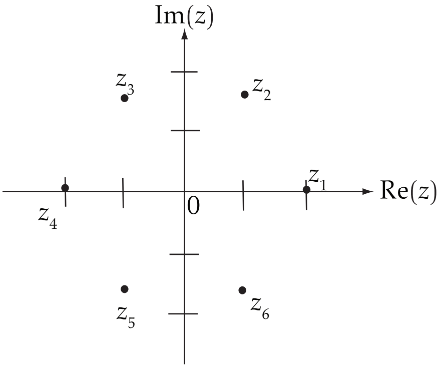


e



Student activity 5.2

a



b for $z = \pm 2$ $\quad |z| = 2$

for $z = 1 \pm i\sqrt{3}$ $\quad |z| = 2$

for $z = -1 \pm i\sqrt{3}$ $\quad |z| = 2$

c z_1 angle for x-axis 0

z_2 angle for x-axis $\frac{\pi}{3}$

z_3 angle for x-axis $\frac{2\pi}{3}$

z_4 angle for x-axis π

z_5 angle for x-axis $\frac{4\pi}{3}$

z_6 angle for x-axis $\frac{5\pi}{3}$

d All point are the same distance from the origin. The angle between each pair of consecutive complex vectors is $\frac{\pi}{3}$.

Student activity 6.1

a $f(z) = u + iv$

$u = \frac{x}{x^2 + y^2} \qquad v = \frac{-y}{x^2 + y^2}$

b $u = x(x^2 - 3y^2) \qquad v = y(3x^2 - y^2)$

$u = x^2 + 2x - y^2 + 1 \qquad v = 2(x + 1)y$

Student activity 6.5

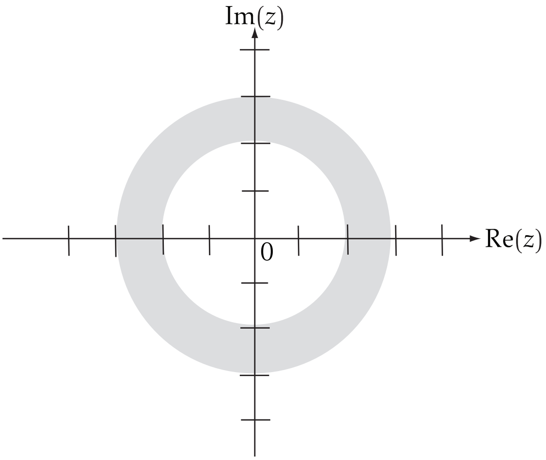
a $|z - 1 + i| \leq 2$

b $\frac{\pi}{2} \leq \arg(z) \leq \pi$

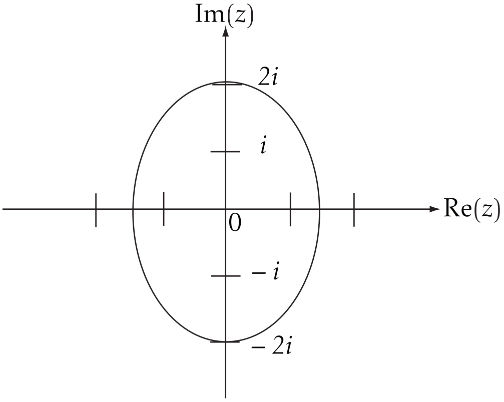
$|z - 1| + |z + 1| < 3$

Student activity 6.6

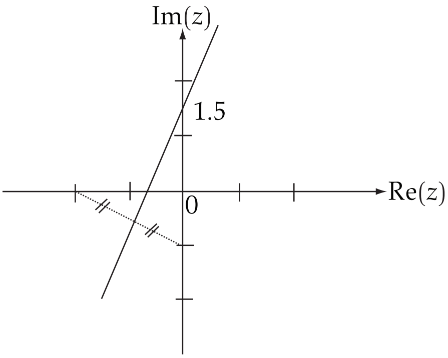
a



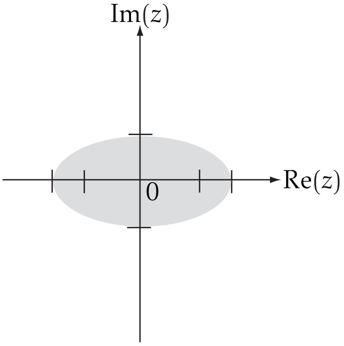
b



c



d



Student activity 8.1

There are a number of approaches to establish these results. Students should be encouraged to explore them all rather than using a single strategy.

For example consider the property $\underset{\sim}{v}.\underset{\sim}{v} = |\underset{\sim}{v}|^2$

a Using the definition of the scalar product

$\underset{\sim}{a}.\underset{\sim}{b} = |\underset{\sim}{a}||\underset{\sim}{b}|\cos(\theta)$

$\underset{\sim}{v}.\underset{\sim}{v} = |\underset{\sim}{v}||\underset{\sim}{v}|\cos(\theta)$

Because $\underset{\sim}{v}$ would be parallel to itself

$\theta = 0^\circ$

$\cos(0) = 1$

$\Rightarrow \underset{\sim}{v}.\underset{\sim}{v} = |\underset{\sim}{v}||\underset{\sim}{v}| = |\underset{\sim}{v}|^2$

Or

Let $\underset{\sim}{v} = x\underset{\sim}{i} + y\underset{\sim}{j} + z\underset{\sim}{k}$

LHS $= \underset{\sim}{v}.\underset{\sim}{v} = x^2 + y^2 + z^2$

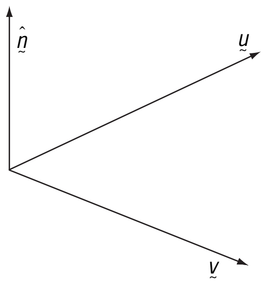
$|\underset{\sim}{v}| = \sqrt{x^2 + y^2 + z^2}$

RHS $= |\underset{\sim}{v}| = x^2 + y^2 + z^2$

LHS = RHS as required

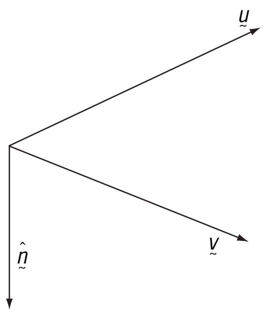
While both approaches will work, many students will feel more comfortable using the second method.

g $\underset{\sim}{v} \times \underset{\sim}{u} = -\underset{\sim}{u} \times \underset{\sim}{v}$



Using the right-hand rule, the vector $\hat{\underset{\sim}{n}}$ would be upwards

$\underset{\sim}{u} \times \underset{\sim}{v} = |\underset{\sim}{u}||\underset{\sim}{v}|\cos(\theta)\hat{\underset{\sim}{n}}$



Using the right-hand rule, the vector $\hat{\underset{\sim}{n}}$ would be downwards

Therefore $\underset{\sim}{v} \times \underset{\sim}{u} = -\underset{\sim}{u} \times \underset{\sim}{v}$

Student activity 9.1

1 a $2\pi \times \frac{60}{\pi} = 120$ seconds or 2 minutes

b $\left(15\cos\left(\frac{45\pi}{60}\right), 15\sin\left(\frac{45\pi}{60}\right)\right) = \left(15\cos\left(\frac{3\pi}{4}\right), 15\sin\left(\frac{3\pi}{4}\right)\right)$

$= \left(15 \times -\frac{1}{\sqrt{2}}, 15 \times \frac{1}{\sqrt{2}}\right)$

$= \left(-\frac{15\sqrt{2}}{2}, \frac{15\sqrt{2}}{2}\right)$

c $x = 15\cos\left(\frac{\pi t}{60}\right)$, $y = 15\sin\left(\frac{\pi t}{60}\right)$

So: $x^2 + y^2 = 15^2\cos^2\left(\frac{\pi t}{60}\right) + 15^2\sin^2\left(\frac{\pi t}{60}\right)$

$x^2 + y^2 = 15^2$

i.e. $x^2 + y^2 = 225$

2 a i $x = 15\cos\left(\frac{\pi t}{60}\right)$ so that

$\frac{dx}{dt} = -15 \times \frac{\pi}{60}\sin\left(\frac{\pi t}{60}\right)$

$= -\frac{\pi}{4}\sin\left(\frac{\pi t}{60}\right)$

$= -\frac{\pi}{4}\sin\left(\frac{3\pi}{4}\right)$ at $t = 45$

$= -\frac{\pi\sqrt{2}}{8}$ (m/s) $[= -0.56]$

a ii $\frac{dy}{dt} = \frac{\pi}{4}\cos\left(\frac{\pi t}{60}\right)$

$= \frac{\pi}{4}\cos\left(\frac{3\pi}{4}\right)$ at $t = 45$

$= -\frac{\pi\sqrt{2}}{8}$ (m/s) $[= -0.56]$

b i If x_d is the x-coordinate of the ball, then from 2ai,

$\frac{dx_d}{dT_d} = -\frac{\pi\sqrt{2}}{8}$

So: $x_d = -\frac{\pi\sqrt{2}}{8}T_d + c$

From 1b, at $T_d = 0$, $x_d = -\frac{15\sqrt{2}}{2}$, hence

$x_d = -\frac{15\sqrt{2}}{2} - \frac{\pi\sqrt{2}}{8}T_d$

(Note: students may use constant acceleration formula.)

b ii If y_d is the y-coordinate of the ball, then

$$\frac{d^2 y_d}{dT_d^2} = -g$$

So: $\frac{dy_d}{dT_d} = -\frac{\pi\sqrt{2}}{8} - gT_d$ (using 2aii at $T_d = 0$)

Then: $y_d = \frac{15\sqrt{2}}{2} - \frac{\pi\sqrt{2}}{8}T_d - \frac{1}{2}gT_d^2$ (by 1b at $T_d = 0$)

(Note: students may use constant acceleration formula.)

3 a i $u\cos(\theta)$ (horizontal acceleration is zero)

a iii $\frac{d^2 y}{dT_b^2} = -g$

So: $\frac{dy}{dT_b} = -gt + c$

At $T_b = 0$, vertical component of velocity is $u\sin(\theta)$, hence the vertical velocity of the ball is $u\sin(\theta) - gT_b$

(Note: students may use constant acceleration formula.)

b i From 3ai:

$$\frac{dx}{dT_b} = u\cos(\theta)$$

So: $x = u\cos(\theta)T_b + c$

At $T_b = 0$, $x = -16$ hence

$$x = -16 + (u\cos(\theta))T_b$$

(Note: students may use constant acceleration formula.)

b ii From 3aii:

$$\frac{dy}{dT_b} = u\sin(\theta) - gT_b$$

So: $y = (u\sin(\theta))T_b - \frac{1}{2}gT_b^2 + c$

At $T_b = 0$, $y = -16$, hence

$$y = -16 + (u\sin(\theta))T_b - \frac{1}{2}gT_b^2$$

(Note: students may use constant acceleration formula.)

c From bi we have $T_b = \frac{x+16}{u\cos(\theta)}$

Substituting for T_b in the expression in bii gives:

$$y = -16 + (u\sin(\theta))\left(\frac{x+16}{u\cos(\theta)}\right) - \frac{1}{2}g\left(\frac{x+16}{u\cos(\theta)}\right)^2$$

$$= -16 + (x+16)\tan(\theta) - \frac{1}{2}g\left(\frac{x+16}{u\cos(\theta)}\right)^2$$

4 a Ball reaches top of its path when vertical velocity is zero:

$u\sin(\theta) - gT_b = 0$

i.e. $T_b = \frac{u\sin(\theta)}{g}$

Also, the ball is at the point $(-15,0)$ at this value of T_b.

So: $\dfrac{-15+16}{u\cos(\theta)} = \dfrac{u\sin(\theta)}{g}$ (equating expressions for T_b)

i.e. $u^2\sin(\theta)\cos(\theta) = g$ (1)

Substituting into the cartesian equation of the ball's paths:

$$0 = -16 + (-15+16)\tan(\theta) - \frac{1}{2}g\left(\frac{-15+16}{u\cos(\theta)}\right)^2$$

$$= -16 + \tan(\theta) - \frac{1}{2}g\frac{1}{u^2\cos^2(\theta)} \quad (2)$$

From (1), $u^2 = \dfrac{g}{\sin(\theta)\cos(\theta)}$

Substituting into (2) gives:

$$0 = -16 + \tan(\theta)\frac{1}{2}g\frac{1}{\cos^2(\theta)}\frac{\sin(\theta)\cos(\theta)}{g}$$

$$= -16 + \tan(\theta) - \frac{1}{2}\tan(\theta) = -16 + \frac{1}{2}\tan(\theta)$$

So: $\tan(\theta) = 32$, giving $\theta = 88.21^\circ$

Then from (1): $u^2 = 313.906$, giving $u = 17.72$ (m/s).

b Ball reaches $(-15,0)$ at time $T_b = \dfrac{u\sin(\theta)}{g} = 1.807$ (s)

Peter takes 60 s to reach $(-15,0)$ from his initial position, he was at the point where $t = 60 - 1.807 = 58.193$ when the ball was released.

So Peter's position was:

$$\left(15\cos\left(\frac{58.193\pi}{60}\right), 15\sin\left(\frac{58.193\pi}{60}\right)\right) = (-14.93, 1.42)$$

5 As in 4a, substitution into the cartesian equation gives

$$0 = -16 + (-15+16)\tan(\theta) - \frac{1}{2}g\left(\frac{-15+16}{u\cos(\theta)}\right)^2$$

So: $-16 + \tan(\theta) - \dfrac{1}{2}g\dfrac{1}{u^2\cos^2(\theta)} = 0$

$$\frac{1}{2}g\frac{1}{u^2\cos^2(\theta)} = -16 + \tan(\theta)$$

Hence: $u^2 = \dfrac{g}{2\cos^2(\theta)(\tan(\theta)-16)}$

i.e. $u = \sqrt{\dfrac{g}{2\cos^2(\theta)(\tan(\theta)-16)}}$

for $\tan(\theta) > 16$ and $\theta < 90$ (approx. $86.24 < \theta < 90$)

REFERENCES AND FURTHER READING

Adler, I 1966, *A new look at geometry*, John Day, New York.
Allendoerfer, CB & Oakley, CC 1963, *Principles of mathematics*, McGraw-Hill, New York.
Bell, ET 1937, *Men of mathematics*, Simon & Schuster, New York.
Boole, G 1847, *The mathematical analysis of logic*, Macmillan, London.
Boyer, CB 1989, *A history of mathematics*, 2nd edn, UC Merzback (revn ed.), Wiley, New York.
Bressoud, DM & Wagon, S 2000, *A course in computational number theory*, Springer-Verlag, London.
Crossley, JN 1987, *The emergence of number*, World Scientific, Singapore.
Deakin, MAB 1996, 'The origins of our number words', *Australian Mathematical Society Gazette*, 23, 2, 50–66.
Dedron, P & Itard, J 1973, *Mathematics and mathematicians*, vols 1 & 2, JV Field (trans.), Open University Press, Milton Keynes.
Kline, M 1972, *Mathematical thought from ancient to modern times*, Oxford University Press, New York.
Martin, GE 1982, *Transformation geometry: An introduction to symmetry*, Springer-Verlag, New York.
Mazur, B 2003, *Imagining numbers: Particularly the square root of minus fifteen*, Penguin, London.
Needham, T 1997, *Visual complex analysis*, Oxford University Press, New York.
Plato, 1871, *The Meno*, B Jowett (trans.), web edition at http://www.mdx.ac.uk/www/study/xplameno.htm
Skemp, R 1989, *Structured activity for primary mathematics: How to enjoy real mathematics*, vols 1 & 2, Routledge, London.
Sobel, D 1995, *Longitude*, Walker, New York.
Stillwell, J 2005, *The four pillars of geometry*, Springer-Verlag, New York.
Struik, DJ 1948, *A concise history of mathematics*, Dover, New York.
Wason, D 2003, *Battlefield detectives*, Granada Media, London.