MAT1102

Algebra & Calculus I

Faculty of Sciences

Study Book

Written by

Associate Professor Chris Harman and Patricia Cretchley
Faculty of Sciences
The University of Southern Queensland

Published by

The University of Southern Queensland
Toowoomba Qld 4350
Australia

http://www.usq.edu.au

Produced using LaTeX in the USQ style.

Preface

Welcome to Algebra & Calculus I. This Study Book is intended to direct your path through the prescribed textbooks, readings, exercises and other activities of this course. It should be used in conjunction with the Introductory Book, which offers a Study Schedule of the sections to be covered week by week.

There are two modules, studied in parallel:

Module 1: Algebra
Module 2: Calculus.

Besides this Study Book, the prescribed texts are

ELEMENTARY LINEAR ALGEBRA, Fifth Edition (2004), by Larson, Edwards & Falvo, published by Houghton Mifflin; and

CALCULUS: Concepts & Contexts, Third Edition (2004 or 2005), by James Stewart, published by Brooks/Cole.
Note: If you have studied this course previously with the previous Calculus textbook by Hughes-Hallett (3rd edition), you will be able to use it together with your previous Study Book. The course website will guide you in the slightly different study order for the material.

In this course, you will be required to apply concepts and processes to both new and familiar situations, and to communicate your methods and results using the appropriate terminology and notation.

Reading the mathematics in your textbooks is just a start, and for most students the assignments will offer inadequate practice. Aim at mastering a broad selection of the suggested exercises, to build experience and gain confidence.

Table of Contents

I Algebra **1**

0 Vectors in the Plane and in Space **7**

0.1 Vectors in the Plane 9

0.2 The Scalar Product and Projections in R^2 11

0.3 Vectors in Space 17

0.4 The Cross Product of Two Vectors 18

0.5 Lines and Planes in Space 20

0.6 Summary: Vectors 28

1 Systems of Linear Equations **29**

1.1 Introduction to Systems of Linear Equations 30

1.2 Gaussian Elimination and Gauss-Jordan Elimination 34

1.3 Application of Systems of Linear Equations 36

1.4 Chapter Summary: Systems of Linear Equations. 36

2 Matrices **37**

2.1 Operations with Matrices 38

2.2 Properties of Matrix Operations 39
2.3 The Inverse of a Matrix 40
2.4 Elementary Matrices 41
2.5 Applications of Matrix Operations 41
2.6 Summary: Matrices 42

3 Determinants 43

3.1 The Determinant of a Matrix 44
3.2 Evaluation of a Determinant using Elementary Operations . . 46
3.3 Properties of Determinants 47
3.4 Chapter Summary on Determinants. 48

4 Complex Numbers 49

4.1 Complex Numbers: Definition 50
4.2 Conjugates and Division of Complex Numbers 51
4.3 Polar Form and de Moivre's Theorem 52
4.4 Euler's Form of a Complex Number 54
4.5 Solving Equations 60
4.6 Functions of Complex Variable z 66
4.7 Summary: Complex Numbers 71

II CALCULUS 73

1 Functions and Models 77

1.1 Four Ways to Represent a Function 78
1.2 Mathematical Models: a Catalog of Essential Functions . . . 79
1.3 New Functions From Old Functions. 81
1.4 Graphing Calculators and Computers 82

1.5 Exponential Functions 82
1.6 Inverse Functions and Logarithms 83

2 Limits and Derivatives **85**
2.1 The Tangent and Velocity Problems 86
2.2 The Limit of a Function 86
2.3 Calculating Limits Using the Limit Laws 87
2.4 Continuity 88
2.5 Limits Involving Infinity 89
2.6 Tangents, Velocities, and Other Rates of Change 89
2.7 Derivatives 90
2.8 The Derivative as a Function 91
2.9 What does f' say about f? 93

3 Differentiation Rules **95**
3.1 Derivatives of Polynomials and Exponential Functions 96
3.2 The Product and Quotient Rules 97
3.3 Rates of Change in the Natural and Social Sciences 98
3.4 Derivatives of Trigonometric Functions 99
3.5 The Chain Rule 99
3.6 Implicit Differentiation 100
3.7 Derivatives of Logarithmic Functions 101

4 Applications of Differentiation **103**
4.1 Related Rates 104
4.2 Maximum and Minimum Values 104
4.3 Derivatives and the Shapes of Curves 105
4.4 Graphing with Calculus and Calculators 106

4.5 Indeterminate Forms and l'Hospital's Rule 106
4.6 Optimisation Problems 106
4.7 Applications to Business and Economics 107
4.8 Newton's Method 107
4.9 Antiderivatives 107

5 Integrals 109

5.1 Areas and Distances 110
5.2 The Definite Integral 111
5.3 Evaluating Definite Integrals 112
5.4 The Fundamental Theorem of Calculus 113
5.5 The Substitution Rule 114
5.6 Integration by Parts 115
5.7 Additional Techniques of Integration 116
5.8 Integration Using Tables and Computer Algebra Systems . . 117
5.9 Approximate Integration 117

6 Applications of Integration 119

6.1 More about Areas 120
6.2 Volumes 120
6.3 Arc Length 121
6.4 Average Value of a Function 121
6.5 Applications to Physics and Engineering 121

III Appendix A and B: Readings for Module I 123

Module I

Algebra

Overview

This module is a collection of topics on *Vectors, Matrices* and *Complex Numbers.* The collective title "Algebra" should not suggest that the treatment is largely algebraic, but that each of these has an associated algebra, ie these quantities can be added, subtracted, and multiplied in ways that proves fruitful for applications. Although these quantities are explored in separate sections, their strong geometric and algebraic links will become clear as you progress through the sections. At the same time, their individual strengths for different types of applications should also become clearer.

Vectors were developed to represent and study physical phenomena and motion in real-world applications. We look at their geometric properties in two and three dimensions, establish rules for vector algebra, and use vectors to characterise lines and planes in space.

Matrices were developed to manage the storage and manipulation of large sets of data and large systems of linear equations which arise with astonishing frequency in scientific, industrial and economic applications. We look also at the algebra of matrices, their addition and subtraction, and define matrix multiplication in such a way that it proves useful for many applications, particularly to represent and solve systems of linear equations. We look at examples of such systems, and consider ways to find solutions.

Complex numbers are a fascinating finale. A brief introduction to their basic properties completes this module. While they arose historically out of attempts to create square roots of negative numbers, their geometric and algebraic properties yield surprisingly practical applications, and complex numbers and their functions are commonly used to describe and analyse a host of real-world phenomena.

Module Objectives

Mathematics is not a collection of isolated ideas and topics. Find and understand the links between the concepts you study. Read with pencil in hand. Confirm the steps numerically and/or algebraically, and draw diagrams where appropriate. Visualisation is a powerful aid. Use technology where it is helpful and encouraged. The assignment tasks and support offered in this course will help you to develop that skill.

At the end of each Chapter you should

- be familiar with all the concepts covered in the Readings
- be able to relate the algebra and geometry where appropriate
- have mastered exercises like those in your Activities, and have practised using the terminology and notation
- be able to communicate your work well, with correct use of appropriate language and symbols
- be able to apply the concepts and processes in a variety of contexts, even unfamiliar ones

At the end of Module 1, Algebra, you should

- see the links between vectors, matrices and complex numbers, and also be aware of the differences between them which make each valuable
- distinguish between their notations, and use them appropriately

At the end of this Course, you should be able to

- use these concepts to support your study in other areas and disciplines, and in Algebra & Calculus II.

Chapter 0

Vectors in the Plane and in Space

You need a good understanding of vectors before you study Chapters 1, 2 and 3 of your prescribed textbook (Elementary Linear Algebra, Edition 4 by Larson & Edwards, or Edition 5 by Larson, Edwards & Falvo). That is why we include this preliminary section titled Chapter 0.

Note that there are readings for this chapter in Appendix A at the back of this Study Book. Appendix A is separated by a yellow page.

The readings in Appendix A are from *Elementary Linear Algebra*, 4th Edition, by Grossman, pp 154–207. This text offers quite full worked solutions for the odd numbers.

In this Study Book Algebra Module, there will also be some readings from your set Calculus textbook *Calculus: Concepts and Contexts*, 3nd Edition, by Stewart. It offers a concise style and more on vectors used in various contexts.

Many quantities are fully characterised once their magnitude is known: eg mass, length, volume, speed. They are known as *scalar* quantities.

Others are only fully specified when we give their direction as well: eg the motion of a boat (speed and direction), wind velocity (strength and direction), forces in a force-field, fluid flow. *Vectors* are quantities that require more than one number for their specification.

There are many kinds of vectors. In this course we study *geometric vectors in 2 and 3 dimensions*. We use their *magnitude* and *direction* to represent movement or motion in the plane or in space, and we use their algebraic and geometric properties to describe lines and planes in 3-dimensional space. In Algebra & Calculus II, we generalise these definitions to vectors of greater dimension.

This introduction to geometric vectors includes:

- geometric representation as an arrow
- coordinate representation as (x, y) or (x, y, z)
- addition and subtraction, to find resultants
- use of vectors to describe motion
- dot product to evaluate lengths and angles
- dot product as a test for perpendicularity
- cross product to create a perpendicular vector
- cross product to describe some areas and volumes
- use of vectors to find equations of lines and planes in space

0.1 Vectors in the Plane

We introduce important basics using geometric vectors in the plane, R^2.

> **Reading**
>
> Appendix A: study section 3.1, from *Elementary Linear Algebra*, 4th Edition, by Grossman. Be sure Table 3.1 makes sense.
> Also read Stewart: sections 9.1 and 9.2. This will be particularly helpful when you reach 3-dimensional vectors in Section 0.3 of this Module.

You may use any of the following notations for vectors: each has its strengths: $(a, b) = \langle a, b\rangle = a\vec{i} + b\vec{j}$.

Be sure you have mastered the following:

- representing a vector *geometrically* as an arrow or directed line segment, and *algebraically* as an ordered pair in round or pointed brackets;
- finding its *magnitude* (length) from components, using Pythagoras;
- finding its *direction angle*: it helps to sketch the vector so as to choose the correct quadrant angle;
- finding *scalar multiples* geometrically and algebraically;
- *adding or subtracting* algebraically, and geometrically using a triangle or parallelogram;
- defining $\vec{i}$ and $\vec{j}$ and expressing vector (a, b) as $a\vec{i} + b\vec{j}$.

Note that if $\vec{u}$ and $\vec{v}$ form adjacent sides of a parallelogram, then their sum and difference form the two *diagonals*. Show this on a sketch.

Be sure you can distinguish between the directions of $\vec{u} - \vec{v}$ and $\vec{v} - \vec{u}$.

Note: Grossman and Stewart use single bars to denote the magnitude of $\vec{v}$. Some texts use double bars $||\vec{v}||$ to emphasise it is *not the absolute value* of a scalar quantity. Use the notation you favour, but it is important to recognise and be familiar with both.

Activity

Note also the following rules for vector algebra. Check geometrically that each rule seems reasonable. Try some examples to get a feeling for why the rule works, but you must use general vectors to prove the rules algebraically: ie of the form $\vec{a} = (a_1, a_2), \vec{b} = (b_1, b_2)$, in R^2, say.

For any vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$, and scalars (numbers) α and β:

- $\vec{a} + \vec{O} = \vec{a}$
- $0\vec{a} = \vec{O}$ (note 0 is a scalar and $\vec{O}$ a vector, so $\vec{O} = (0, 0)$ in R^2);
- $\vec{a} + \vec{b} = \vec{b} + \vec{a}$
- $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$
- $\alpha(\vec{a} + \vec{b}) = \alpha\vec{a} + \alpha\vec{b}$
- $(\alpha + \beta)\vec{a} = \alpha\vec{a} + \beta\vec{a}$
- $(\alpha\beta)\vec{a} = \alpha(\beta\vec{a})$

To make any vector $\vec{v}$ into a unit vector, note that we can simply divide it (ie each of its components) by its own length: $\frac{\vec{v}}{|v|}$ or $\frac{\vec{v}}{||v||}$.

Activity

Do Grossman, Problems 3.1 of Appendix A, odd numbers 1-35 (31 optional), and as many others as you can manage. Note that Grossman gives quite full worked solutions for the odd numbers. These are a valuable aid.
Use examples and exercises from Stewart's Section 9.2 to support your study, if you find that helps.

0.2 The Scalar Product and Projections in R^2

Addition and subtraction of vectors are done *componentwise*, and so is multiplication by a scalar. But we will not define the product of vectors (u_1, u_2) and (v_1, v_2) componentwise as (u_1v_1, u_2v_2) here, because this definition turns out to be less useful in the applications we need now, than the so-called **dot** and **cross** products.

We study **cross or vector product** in Section 0.4.

The **dot or scalar** product of two vectors,

$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2$$

is so-called because it gives a *scalar* from two vectors.

Note the most important property:

$$\vec{u} \cdot \vec{v} = |\vec{u}|\,|\vec{v}| \cos\phi.$$

It is used

- for finding the angle between $\vec{u}$ and $\vec{v}$, and
- to give a test for perpendicularity: two vectors are perpendicular if and only if their dot product is 0.

> **Reading**
>
> Study Appendix A, Grossman, Section 3.2 in two parts: first scalar (or dot) products, pp 164–mid 168. Continue with projections when you have mastered dot products.

Theorem 1 referred to on p 166 of Appendix A is given below. Its parts can be proved algebraically by expressing each side in terms of components.

Theorem 1: If $\vec{a}, \vec{b}$ and $\vec{c}$ are any vectors, and α is any scalar (number), then

1. $\vec{a} \cdot \vec{O} = 0$
2. $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$
3. $\vec{a} \cdot (b \vec{+} c) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$
4. $(\alpha\vec{a}) \cdot \vec{b} = \alpha(\vec{a} \cdot \vec{b})$

Be sure you have mastered the following:

- calculating dot product from the components of the vectors:
$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2;$$
- using the formula $\vec{u} \cdot \vec{v} = |\vec{u}|\,|\vec{v}| \cos\phi$ to find the angle between two vectors:
$$\cos\phi = \frac{\vec{u} \cdot \vec{v}}{|\vec{u}|\,|\vec{v}|};$$
- *dotting a vector with itself* (to put it very loosely) to find the square of its own length;
- how to show two vectors are parallel: they must be multiples of each other, like $(2, -1)$ and $(6, -3)$ (Th 3);

Important:
Theorem 4 gives us the *dot product test for perpendicularity*: non-zero vectors $\vec{u}$ and $\vec{v}$ are perpendicular if and only if $\vec{u} \cdot \vec{v} = 0$.

Note the test for parallelness in the Activity below: parallel vectors are scalar multiples of each other.

Because the sign of the dot product is determined by the cosine of the angle, it gives the relative directions of the vectors:

The angle between the vectors is

- acute (ie the vectors pull together, generally) if and only if their dot product is positive;
- obtuse (ie the vectors pull away from each other) if and only if their dot product is negative;

- a right angle if and only if their dot product is 0.

Activity

Do Appendix A, Problems 3.2, odd numbers,1-19, to ensure mastery.
Note Q 35 which gives a simple test for parallelness:
two non-zero vectors are parallel iff (if and only if) one is a scalar multiple of the other:
for example $(-2,3)$ and $(6,-9)$ are parallel
because $(6,-9) = -3(-2,3)$.

Sketch to check. Note that the negative multiple tells us that they point in opposite directions.

Using Vectors to Describe Motion:

Activity

Stewart: study section 9.2 and do Exercises 1, 3, 5, 23, 24, 25.

Note on Exercise 24 above:

Draw the two vectors for the wind and the plane. To add them, attach the tail of the plane's vector to the head of the wind's vector. The resultant course is the third side of the triangle, where the base angle is 45+60=105 degrees. You can then use the cosine rule to find this third side. Then you can use the sine rule to find the angles of the triangle.

Components and Projections

Reading

Study components and projections.
Use Stewart Section 9.3, pp 654–655 , or the notes below, to replace Grossman, Sec 3.2, lower p 168.
Definition 4, Grossman p 169, gives a summary of the important results. Study Stewart's Example 7, p 655, and Grossman's Ex 5 & 6, p 170.

Notes replacing lower p 168 in Appendix A:

We often want to know how much effect a vector $\vec{u}$ has in some direction other than its own, say that of vector $\vec{v}$; eg how much effect a force vector, like wind or current, has on the movement of an object in another direction. For examples, see The Dot Product section of the Chapter 9 in your Calculus text, Stewart: section 9.3.

If we start $\vec{u}$ and $\vec{v}$ at the same point, we can find how far $\vec{u}$ extends in the direction of $\vec{v}$ by dropping a perpendicular from the tip of $\vec{u}$ onto the direction given by $\vec{v}$. See Figure 3.15(a), p 169, Appendix A.

Note how $\vec{u}$ is then the sum of the two perpendicular (orthogonal) vectors that form the sides of the right-angle in the triangle, ie $\vec{u} = proj_v\vec{u} + \vec{w}$.

We say $\vec{u}$ can be *decomposed (or resolved) into the sum of those two vector components*, one parallel to $\vec{v}$ and one perpendicular to $\vec{v}$.

Scalar component:

If θ is the angle between $\vec{u}$ and $\vec{v}$, then the definition of cosine tells us that the vector $proj_v\vec{u}$ has length $|\vec{u}|\cos\theta$.

But $\cos\theta$ may be positive or negative, so $|\vec{u}|\cos\theta$ is a *signed* length expressing how far $\vec{u}$ extends in direction $\vec{v}$.

We call $|\vec{u}|\cos\theta$ the *scalar component* of $\vec{u}$ in the direction of $\vec{v}$.

Note that dividing by $|\vec{v}|$ in $\vec{u}\cdot\vec{v} = |\vec{u}||\vec{v}|\cos\theta$, gives

$$\text{Scalar component } = |\vec{u}|\cos\theta = \frac{\vec{u}\cdot\vec{v}}{|\vec{v}|}$$

where the RHS is easier to calculate. See Equation (5),Appendix A p 169.

Vector Projection:

If we write this component as a *vector*, rather, then we call it the *projection* of $\vec{u}$ in the direction of $\vec{v}$, or the *orthogonal projection* of $\vec{u}$ on $\vec{v}$, written $proj_v\vec{u}$.

This projection vector simply expresses the scalar component, $\frac{\vec{u}\cdot\vec{v}}{|\vec{v}|}$ or $|\vec{u}|\cos\theta$, as a vector in the direction of $\vec{v}$. How can we do that?

Think of a simpler task: to turn scalar -5 into a vector in the direction of the x-axis, we multiply it by a *unit* vector in that direction, $\vec{i}$, to get $-5\vec{i}$. Similarly, a vector 7 units long in the direction of the positive y-axis is given by the vector $7\vec{j}$.

To turn a signed scalar length into a vector in a specific direction, all we need to do is multiply it by a unit vector in the required direction.

Hence scalar component $\frac{\vec{u}\cdot\vec{v}}{|\vec{v}|}$ must be multiplied by a *unit vector* in the direction of $\vec{v}$, ie $\frac{\vec{v}}{|\vec{v}|}$. We conclude that

$$proj_v\vec{u} = |\vec{u}|\cos\theta\frac{\vec{v}}{|\vec{v}|} = \left(\frac{\vec{u}\cdot\vec{v}}{|\vec{v}|}\right)\frac{\vec{v}}{|\vec{v}|} = \left(\frac{\vec{u}\cdot\vec{v}}{|\vec{v}|^2}\right)\vec{v}$$

which is easier to calculate. See Equation (4), p 169.

Note: We can use this to *resolve* or *decompose* $\vec{u}$ into the sum of two vector components, one in the direction of $\vec{v}$, and the other perpendicular to that, ie $proj_v\vec{u}$ and $\vec{w}$ as shown in Figure 3.15, App A p 169.

To decompose, we find $proj_v \vec{u}$ first, as above, then find $\vec{w}$ by subtracting it from $\vec{u}$ (because $\vec{u} = proj_v \vec{u} + \vec{w}$.)

Activity

You may find it helpful now to study Projections in Stewart: section 9.3.
Then do Appendix A, Grossman: problems 3.2, odd numbers 1-19, 21-27, 33, 35 and 39. Master more, if you have time.

0.3 Vectors in Space

In this section, 2-dimensional geometric vectors in the plane, R^2, are extended to 3-dimensional space, R^3, by the inclusion of a third component in the z-direction. The rules and definitions already established for R^2 extend easily, with a little checking.

Reading

Read Appendix A, Grossman, Section 3.3.
Take the emphasis off p 177, 178, direction cosines. Example 5 is optional.

Also read Stewart Sec 9.1, 637–640.

Activity

Do Appendix A, Grossman problems 3.3, odd numbers 1-9, 21, 23, odds 27-39. Do as many others as you can manage.

0.4 The Cross Product of Two Vectors

The **dot product** of any vectors $\vec{u}$ and $\vec{v}$ in R^2 or R^3 gave a formula for the cosine of the angle between the vectors. We have been able to use that as a test for perpendicularity, and to make projections easier to calculate.

The **cross or vector product**, $\vec{u} \times \vec{v}$ defined at the bottom of p 183, Appendix A, is another very useful product of two vectors $\vec{u}$ and $\vec{v}$ in R^3:

cross product is useful because it gives a *vector perpendicular to both $\vec{u}$ and $\vec{v}$*.

> **Reading**
>
> Study Appendix A Section 3.4. Omit mid-187 to mid 188. Also omit Theorem 1 and Example 2 if you are not familiar with determinants: we study them in Chapter 3.
> Use Sec 9.4 from Stewart, pp 657-663, as back-up.

Finding cross products:

Be careful: note that the middle component in the definition on p 183 has the opposite order of subtraction to the other two components, causing a change of sign.

If you know how to expand a determinant Theorem 1 offers a method of calculation. But even if you have never seen a determinant, it is easy to proceed as follows. Simply write the components of the second vector directly beneath the first: eg for the vectors in Example 1, p 184, write

$$\begin{vmatrix} 1 & -1 & 2 \\ 2 & 3 & -4 \end{vmatrix}.$$

To get the first component, block out the first column, then use a cross pattern to multiply the diagonals that remain, ie $\begin{vmatrix} -1 & 2 \\ 3 & -4 \end{vmatrix}$ and subtract them: $(-1)(-4) - (2)(3) = -2$.

Do the same for the other two components, but *for the middle one, remember to subtract in the opposite order!*

Th 2(vi) proves the important useful property of cross product: that $\vec{u} \times \vec{v}$ gives a vector perpendicular (orthogonal) to each of the two vectors crossed.

It does this by showing that $(\vec{u} \times \vec{v}) \cdot \vec{u} = 0$ and $(\vec{u} \times \vec{v}) \cdot \vec{v} = 0$.

Note these important results too:

- $\vec{v} \times \vec{u}$ has *magnitude* $|\vec{u}|\,|\vec{v}| \sin \phi$ (like dot product but with sine instead of cos);
- $\vec{v} \times \vec{u}$ has the direction of the thumb in the right-hand rule;
- $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$ surprisingly;
- the magnitude $|\vec{u} \times \vec{v}|$ also gives the area of a parallelogram;
- the *absolute value* of a *scalar triple product* $|\vec{w} \cdot (\vec{u} \times \vec{v})|$ gives the volume of a parallelepiped.

Activity

Do Problems 3.4, Q 3, 5, 9, 17, 19, 21, 23, 35, 29, 31, 33 37. Do as many others as you can manage.
Note that you can also use cross product to find the **area of a triangle** with adjacent sides $\vec{u}$ and $\vec{v}$: the area is half that of the parallelogram they define.

Activity

Browse through the approach to properties of vectors and their uses in the Vectors Chapter in Stewart: sections 9.1 to 9.4.

Note the sections on dot and cross product, in particular.

0.5 Lines and Planes in Space

We now use dot and cross products to obtain equations for lines and planes in space. Dot product gives a test for perpendicularity. Cross product gives a vector perpendicular to those we cross.

We ask the following two questions now, in 3-dimensional space, R^3:

- What general form of equation describes any given *line*?
- What general form describes any *plane*?

In the plane, R^2, ie two-dimensional space, $ax + by = c$ is the general form of the equation of a line. (Vertical lines of form $x = c$ are not described by the general slope-intercept form, $y = mx + k$, but are included in the more general form, $ax + by = c$, when $b = 0$.)

As we show below, however, extending $ax + by = c$ in R^2 to include the extra variable, $ax + by + cz = d$, does *not* give a *line*in R^3, but describes a *plane*, rather.

> **Reading**
>
> The notes below replace those on the bottom of Grossman, p 193–194 of Appendix A, Section 3.5.
> Stewart: section 9.5, pp 666–673, provides further back-up.

The Equation of a Line:

To specify a particular line, in the plane or in space, it is sufficient to know just one point on it, as long as we know its direction.

Suppose we have a line, firstly in the *plane* for simplicity, with y-intercept 3, and slope 4.

Sketch this line here, and convince yourself that because the slope is 4/1, the vector (1,4) is parallel to it.

Notice how we can travel from the origin to any point (x, y) on the line, using the known point $(0, 3)$ (the y-intercept) and the slope or direction vector, $(1, 4)$:

- Firstly, to reach the line, we can travel along the position vector $(0, 3)$ to get to the only known point on the line.

- Then to move from this point to any other point on the line, say higher, we need to add further travel in the direction of the line, ie in the direction of the slope or direction vector $(1, 4)$.

How far should we travel in the direction of $(1, 4)$? We want to get to any point on the line.

- If we travel exactly 1 such direction vector, we get to the point $(0, 3) + (1, 4)$, ie $(1, 7)$. Confirm this graphically: travel to $(0, 3)$ first, then travel the vector $(1, 4)$ which is parallel to the line.
- If we travel 2 slope vectors $(1, 4)$ instead, after going to $(0, 3)$, we reach the point $(0, 3) + 2(1, 4)$ ie $(2, 11)$.
- Using big positive multiples of the direction vector will take us to points much further up the line.
- To reach points further down the line from $(0, 3)$, we need to travel in the *opposite direction*, ie in the direction of $-(1, 4) = (-1, -4)$. Using ever bigger negative multiples of the direction vector will take us to points further on the other side of $(0, 3)$.

To conclude: as long as we allow t to have all possible positive and negative values, the sum $(0, 3) + t(1, 4)$ will take us to all possible points (x, y) on the line. This means that the line comprises all possible points of the form

$$(x, y) = (0, 3) + t(1, 4), \text{ where t is any real number.}$$

There is nothing special about the particular point and direction vector we used: we can travel to *any fixed point* (x_1, y_1) on the line, then simply add multiples t of *any direction vector* (a, b).

For example, we could have used any multiple of $(1, 4)$ as the direction vector in the example above: that would still give a vector parallel to the line. If we used $(2, 8)$ instead, then to reach a particular point, we would need exactly half the multiple required if we were using direction vector $(1, 4)$. So the line is also defined by the vector equation

$$(x, y) = (0, 3) + s(2, 8), \ s \text{ any real number.}$$

In general, the vector equation of a line in R^2 through given point (x_1, y_1), and parallel to the vector (a, b), has form

$$(x, y) = (x_1, y_1) + t(a, b), \text{ t any real number.}$$

This reads as follows:
for some real value of t, (the coordinates of a variable point on the line) = (the coordinates of a fixed point on the line) + t(a direction for the line).

If P is the fixed point (x_1, y_1) on the line, and R is the variable point (x, y) on the line, then $(x, y) = \vec{OR}$ and $(x_1, y_1) = \vec{OP}$. And if we write the direction vector (a, b) as $\vec{v}$, then the vector equation has concise form:

$$\vec{OR} = \vec{OP} + t(\vec{v}), \text{ t any real number.}$$

This generalises easily to 3-dimensional space: the vector equation of the line through the point $P(x_1, y_1, z_1)$ and parallel to the vector $\vec{v} = (a, b, c)$, is

$$\vec{OR} = \vec{OP} + t\vec{v}, \text{ t any real number; or}$$

$$(x, y, z) = (x_1, y_1, z_1) + t(a, b, c), \text{ t any real number.}$$

Example:

The equation of the line that passes through point $(1, 4, 3,)$ in a direction parallel to $(3, -1, 2)$, is $(x, y, z) = (1, 4, 3) + t(3, -1, 2)$, t any real number.

To find points on the line, we simply put in values for t:
$t = 0$ gives the point $(x, y, z) = (1, 4, 3)$.
$t = 1$ gives the point $(x.y.z) = (1, 4, 3) + (3, -1, 2) = (4, 3, 5)$.
$t = -1$ gives $(x, y, z) = (1, 4, 3) - (3, -1, 2) = (-2, 5, 1)$, etc.

Note that we often have to deduce the direction vector from other given information about slope, parallelness, or orthogonality. It is *any vector* $\vec{v} = (a, b, c)$ parallel to the line.

For example, if we are given two points P and Q on the line, we can use the joining vector $\vec{PQ}$ (obtained by subtracting the components) as a direction vector. See Example 1 of Appendix A, p 195, and equation (1) on p 193.

We call t the *parameter*. Different values of the parameter give different points on the line.

- To get the coordinates of a particular point we give t a particular value.
- If we use all possible real numbers for t, we get all the points on the line.
- If we use an interval of values, say $-2 \leq t \leq 3$, we will get the points on a particular part of the line: that line segment that stretches three multiples of $\vec{v}$ from the fixed point P, and 2 multiples of $\vec{v}$ in the opposite direction.

The Three Forms of the Equation of a Line in R^3:

So far, we have discussed **vector form**:

$$(x, y, z) = (x_1, y_1, z_1) + t(a, b, c), \text{ with t any real number.}$$

Parametric form follows by equating the x, y and z components on each side, and separating them. This gives equations like those in (5) on p 194:

$$\begin{aligned} x &= x_1 + ta \\ y &= y_1 + tb \\ z &= z_1 + tc, \end{aligned}$$

with parameter t still any real number.

Symmetric form follows by solving each of the above 3 equations for t, if possible, and then equating them. This gives equations like (6) on p 194:

$$\frac{x - x_1}{a} = \frac{y - y_1}{b} = \frac{z - z_1}{c}.$$

Note that this form requires division by a, b and c. If any of these is 0, the division is done for the others, and the difficult one listed separately: see Appendix A, Example 3, p 195, and Example 5, p 196.

Important: Be sure that in each form of the equation of a line, you can pick out the direction vector (a, b, c) and at least one fixed point (x_1, y_1, z_1) *immediately*. Other points on the line follow by substituting different values for t.

Activity

Study Examples 1, 2 and 3, on p 195 of Grossman in Appendix A. Take the emphasis off Examples 4, 5 and 6. (Also study Stewart: section 9.5 Examples 1, 2 on pp 667, 668)
Then do Problems 3.5, odd numbers 1-3, 9-19, and 23-25. We need vector or parametric forms later, rather than symmetric form, so aim at finding vector form first, and be able to change to parametric form.
Appendix A uses $\vec{i}, \vec{j}$ and $\vec{k}$ notation, but component notation is often quicker and clearer for lines. For example, the equation on line 6 of p 195 can be written simply as $\vec{OR} = (2, -1, 6) + t(1, 2, -8)$.

We now move on to constructing the equation of a *plane*, or indeed, formalising what we mean by a plane in space.

The Equation of a Plane in R^3:

Reading

The notes below are a replacement for most of p 197 of App A, Section 3.5. They are very important.

Study them carefully *before* reading pp 198 - mid 201. Pp 201 and 202 may be omitted, but you may want to come back to Example 10 when you have learned to row-reduce a matrix.

Stewart: section 9.5 pp 669–671, provides further discussion, but you need not study Examples 7, 9 and 10.

Consider a particular plane in space: imagine a flat sheet extending without bound in any direction. We seek a property that distinguishes all the points on that particular plane from other points in space.

A line's direction is characterised by a direction vector parallel to it, because a line has only one such direction. A plane's direction cannot be characterised in that way by a vector parallel to it, however, because many vectors with different directions are parallel to it: consider any vector that joins two points in the plane, for example.

But there is only one direction perpendicular to the plane.

Imagine a table-top or a slanted roof: if we drew vectors perpendicular to the surface they would all be parallel, though some might point in opposite directions.

We say all such vectors are *normal* to the plane, meaning perpendicular, and we call any one of them a **normal** for that plane.

- A plane has many normals: if you have one, any positive or negative multiple will also be normal to that plane.
- Parallel planes have parallel normals. Draw some to see this.
- Planes that are not parallel will have normals that are not parallel. Convince yourself: draw some.

- Planes that are perpendicular, ie orthogonal, have normals that are orthogonal. Convince yourself!

Given a normal, there are infinitely many planes, all parallel, that are perpendicular to it. To distinguish one of them in particular, we need to know at least one point on it. So a plane is fully specified by knowing

- a vector perpendicular or normal to it,
- and a particular point on it.

This information should therefore be enough for us to arrive at an equation that characterises all its points.

Suppose, therefore, that point $P(x_0, y_0, z_0)$ lies on a plane and that vector $\vec{n} = (a, b, c)$ is normal or perpendicular to plane.

Let us show $\vec{n}$ perpendicular to the plane at the given point P, as shown in Figure 3.38 in Grossman Appendix A, p 197. Redraw it here.
(See also in Stewart: section 9.5 p 669 Figure 6, where the point (x_0, y_0, z_0) is labelled P_0, and the variable point (x, y, z) is called P.)

Now suppose $Q(x, y, z)$ is a variable point anywhere on that plane.

Join Q to P. This gives vector $\vec{PQ}$ lying in the plane. Since $\vec{n}$ is normal to the plane, $\vec{n}$ will be perpendicular to $\vec{PQ}$. This is only true if Q is on the plane: if Q is above or below the plane, joining it to P will make an acute or obtuse angle with $\vec{n}$.

So for all points Q on the plane, and no others, $\vec{PQ}$ is perpendicular to $\vec{n}$. Using the dot product test for orthogonality, therefore, we have an equation that characterises points Q on the plane:

$$\vec{PQ} \cdot \vec{n} = 0.$$

Giving $P(x_0, y_0, z_0)$ and $Q(x, y, z)$ their coordinates, we get

$$[(x, y, z) - (x_0, y_0, z_0)] \cdot \vec{n} = 0.$$

Using the distributive rule on the bracket, gives yet another form:

$$(x, y, z) \cdot \vec{n} - (x_0, y_0, z_0) \cdot \vec{n} = 0,$$

and putting the second term on the RHS gives

$$(x, y, z) \cdot \vec{n} = (x_0, y_0, z_0) \cdot \vec{n}.$$

This are all different forms of the equation for the plane, and give us different ways of finding it. Notice in particular the easy method:

(variable point) · (normal) = (fixed point) · (normal).

Vector $\vec{n} = (a, b, c)$, so this gives

$$(x, y, z) \cdot (a, b, c) = (x_0, y_0, z_0) \cdot (a, b, c),$$

and taking dot products gives

$$ax + by + cz = ax_0 + by_0 + cz_0.$$

The LHS cannot be simplified because x, y and z are variables. But the RHS will reduce to an answer, d, say.

So in general the equation of a plane has form $ax + by + cz = d$.

Note also that in this form, the *coefficients* (a, b, c) give a *normal* to the plane.

Example:

1. The plane that passes through the point $(1, 2, -3)$ in such a way that it is normal to the vector $(2, -1, 4)$, has equation

   $$(x, y, z) \cdot (2, -1, 4) = (1, 2, -3) \cdot (2, -1, 4.)$$

   Evaluating the dot products gives

   $$2x - y + 4z = 2 - 2 - 12, \text{ or simply } 2x - y + 4z = -12.$$

   To find points (x, y, z) on this plane, we can substitute values for any two of the variables, and work out the third. For example, substituting easy values $x = 0$ and $y = 0$, we find $4z = -12$, so that $z = -3$. The point $(0, 0, -3)$ therefore lies on that plane.

   Verify by substitution that the point $(1, 1, -\frac{13}{4})$ also lies on the plane, and find another point.

   Also see Example 7, Appendix A, p 198, and the examples labelled i, ii and iii at the bottom of p 198 and the top of p 199.

2. Sometimes the normal is not given, but can be gleaned from other information: eg, if the plane is known to be parallel to another plane with given equation, then one can use the known one's normal (ie its coefficients of x, y and z) as a normal for the required plane;

3. Note that **cross product can be used to create a normal to a plane**, from two vectors that lie in the plane or are parallel to it. For example, given three points P, Q and R in the plane (not all on the same straight line) joining two of them in any direction gives a vector that lies in the plane. Joining any other pair will make another vector in the plane. The cross product of these, say $\vec{PQ} \times \vec{RQ}$, gives a vector perpendicular to those two vectors, and hence also to the plane in which they lie. See example 8, Appendix A, p 199.

Reading

Study any material you have not yet covered in Grossman, pp 198-200 in Appendix A. The material on pp 201-202 is optional and may be omitted, but you may want to return to Example 10 after Sections 1.1 and 1.2.

Be sure that you have mastered the following:

- the general form of the equation of a plane is $ax + by + cz = d$;
- to find points on a plane with a given equation, we can substitute any values we like for two of the variables, and w ork out the corresponding value for the third;
- given an equation of a plane of that form, the coefficients (a, b, c) give a vector *normal* to that plane;
- parallel planes have parallel normals (ie scalar multiples of each other);
- perpendicular planes have normals that are perpendicular (ie have dot product 0).

Activity

Do Grossman, Problems 3.5 in Appendix A, odd numbers 29-49, 50, 51, 53 and 55. Do any more you can manage.

0.6 Summary: Vectors

Activity

Browse through Stewart: sections 9.1 to 9.5 for other perspectives on vectors.
Check Grossman's Summary on p 205-207 in Appendix A. Where something looks unfamiliar, go back and master it. The confidence you gain when the material begins to compact into something manageable is worth the effort.

Do not neglect to do exercises for each section. Assignments do not offer enough breadth or practice!

Chapter 1

Systems of Linear Equations

We know that an equation of form $ax + by = c$ defines a straight line in R^2.

We now too, know that an equation of form $ax + by + cz = d$ describes a plane in R^3, not a line, but we still say it is *linear* in variables x, y, and z, because the variables behave like those of a line in R^2 in that for every step of a fixed size in the x-direction, say, z rises or falls by a constant amount, and similarly, for every step of a constant size in the y-direction, there is a constant amount of rise or fall.

Clearly equations of this kind are significant. Linear equations arise in many areas of science, engineering, technology, business and industry, not just geometrically, but in such a way that they represent demands or constraints on variables it is often the case that many such equations must be satisfied simultaneously. The need to understand, simplify and solve them, is one of the most widely needed areas of application of mathematics.

We now work from your prescribed text, *Elementary Linear Algebra*, by Larson & Edwards, or Larson, Edwards & Falvo. We shall call it Larson *et al.* By way of introduction, read p xiii and xiv of Ed 4 (xv and xvi of Ed 5).

1.1 Introduction to Systems of Linear Equations

Suppose we want to find the points of intersection of a set of lines or planes, as in Example 4 of Larson *et al* Section 1.1. We need to be able to solve the equations simultaneously for the exact coordinates of all possible points of intersection.

Sometimes we can predict the type of intersection simply by using our understanding of the coefficients.

Given the equations of 2 lines in the plane, a quick look at their slopes tells us if they intersect or not. Parallel lines have the slope. Intersecting lines do not. For example, the lines $y = 3x - 2$ and $y = -2x + 5$ must intersect at some point: since their slopes, 3 and -2, respectively, are different.

Given a 3rd line, with a different slope to those of the first two, it may not be easy to tell if it intersects the first two at their common point, or if misses that point, and intersects them each at two other points, thereby giving no common solution to the three equations, simultaneously.

In a similar way, if two planes in 3-dimensional space are not parallel, they must meet along a line: imagine one plane dipping towards the other and cutting through it.

Hence given the equations of two planes, a quick look at their normals tells us if they are parallel or not.

For example, the planes defined by $2x - y + z = 3$ and $-4x + 2y - 2z = 5$ are parallel, because their normals $(2, -1, 1)$ and $(-4, 2, -2)$ are multiples of each other, and hence parallel. They have no intersection.

But the planes $2x - y + z = 3$ and $x - 2y - 4z = 1$ are not parallel, because their normals $(2, -1, 1)$ and $(1, -2 -4)$ are not multiples of each other. Hence these 2 planes meet along a straight line.

If we introduce a third plane, there are many possible ways the three could intersect. Imagine some: they could meet at just 1 point, where the third plane cuts through the line of intersection of the first two. They could meet along a line, like the pages meeting along the spine of a book, or they might coincide, and share all their points. Or they may have no common point of intersection at all: eg they are all parallel but do not coincide; or two are parallel, and the third cuts through them in different places; or they form an A-frame, relative to each other).

Draw some of these possibilities, and see if you can think of others.

> **Reading**
>
> Study Section 1.1 of Larson *et al.* All of the material is important.

Be sure that you have mastered the following:

- be able to tell whether any given equation is *linear* or not
- know the general form of a linear equation
- know what is meant by a *system of linear equations* and its *solutions*, if any
- know that *consistent* means having at least 1 solution, *inconsistent* means having no solutions at all
- be able to express one variable in terms of all the others in a linear equation
- know what is meant by the terms *free* variable and *parameter*
- be able to find *particular solutions* for the other variables by substituting arbitrary values for the parameters

Note that to solve a system of linear equations, we try change the system into a simpler one that retains the same solutions. Note that we use the *elementary row operations* to do that because they do not spoil the solutions.

- Know exactly what the three elementary row operations are.
- Aim at obtaining a *triangular* structure, called *row-echelon* form, or else a *diagonal* form
- If necessary, use back-substitution to find values for the variables

Note also that solving a system of linear equations of any size and in any number of variables, leads to only three possibilities:

- exactly 1 solution,
- infinitely many solutions, or

- no solutions at all; the system is then called *inconsistent*.

Example 4 makes the last point clear for 2 equations in 2 variables. Solutions are the points of intersection, if any, of 2 lines. It is impossible to find a solution which comprises 2 points, or 10 points, say.

We find

- exactly 1 point of intersection (when they cross), or
- infinitely many (when the lines coincide), or
- none at all (when the lines are parallel).

Activity

Confirm that for the intersection of planes in space, we still only get three possibilities: one point, or none, or infinitely many. Draw rough sketches of all the possible ways they can intersect. Start with the intersection of just 2 planes. Then try 3 planes. Examples 7 and 8 illustrate some possibilities. Confirm that there will be

- *exactly 1 point* of intersection, if the third plane cuts through the line of intersection of the first two
- *none at all* if the third plane "misses" the line of intersection of the other two: ie the planes intersect only two at a time as though they frame an A
- *infinitely many* points of intersection along a line if the planes meet like the pages of a book
- *infinitely many* along a plane if the 3 planes all coincide
- *none at all* if all of the planes are parallel
- *none at all* even if only two are parallel, because those two have no points of intersection

Think about the effect of introducing a fourth plane too.

Activity

Section 1.1 Exercises, a selection of odd numbers, 1-25, 35-37, 43.

Also do Ed 4: 51-59, and 73;
or Ed 5: 51-64, 77.

1.2 Gaussian Elimination and Gauss-Jordan Elimination

Reading

Study Section 1.2 of Larson *et al.* We use matrix notation to keep the essential information concise when solving linear equations.

- Note how we can use the *elementary row operations* to reduce a system to simpler ones without altering the solutions. Always record against each row what row operation you did to get it, to help the reader.
- Note the processes referred to as *Gaussian elimination* and *Gauss-Jordan elimination*
- that the former leads to row-echelon form and the latter leads to reduced row-echelon form;
- that using Gauss-Jordan rather than Gaussian elimination may increase the number of row-reduction steps, but saves hand-steps later, by reducing the back-substitution necessary to complete the solving.

These systematic processes contain basic ideas that support the reasoning behind some of our later work, but note that while they lend themselves to machine computation, they are generally far from efficient, and may lead to problems. Refinements like partial or complete pivoting may offer ways of reducing rounding errors, but do not always eliminate the problems, and there is a need for alternative methods appropriate for different types of systems.

Note

- that perfect triangular or diagonal form with 1's on the main diagonal, does not always arise
- how to interpret rows of all 0's as no information at all, and missing rows, as in Examples 8 and 9
- that a row all 0's except the RHS which is non-zero, leads to a contradiction, and that system hence has no solutions: eg $[0\ 0\ 0\ -2]$ really means $0x + 0y + 0z = -2$ which cannot be satisfied by any values of x, y and z

- that a homogeneous system (ie one with RHS's all 0's) always has at least one solution, when all the variables are 0's (called the *trivial solution*) but that it may also have others, infinitely many of the parametric type. So homogeneous systems are always consistent.

Activity

Look again at Example 8. Imagine that the variables were x, y and z. Then the two equations represent planes, and solving them simultaneously will find their point/s of intersection,i f any. Because they are not parallel, we expect them to meet along a line. Show that the solution found there can be written in the form $(x, y, z) = (2, -1, 0) + t(-5, 3, 1)$ and hence deduce that this is the equation of the line of intersection of the planes.
Then do Section 1.2 Exercises, odd numbers 1, 5, 9, 15-21.
Also do Ed 4: 29-43, 59;
or Ed 5: 29-45, 61.

1.3 Application of Systems of Linear Equations

This section covers applications to polynomial curve-fitting, and network analysis. *Here you have a choice!*

Reading

Study Larson Examples 1 to 4, on fitting polynomials to points. Then study at least one type of network: either Example 5, or else electrical circuits as in Examples 6 and 7.

Activity

Do Section 1.3 Exercises, 1, 5, 7, 11 (If you have Ed 5, also do Q15).

Also do two network problems of the type you have chosen: from Ed 4, 19-26 (or Ed 5, 21-28).

1.4 Chapter Summary: Systems of Linear Equations.

Remind yourself of the Module objectives listed immediately before Chapter 0.

Activity

Check through the summary at the end of Section 1. You can find another on *Systems of Linear Equations* in the coloured pages of the Study Guide Appendix A.

Attempt some of the Review Exercises in Larson *et al*, and look at the Projects.

Chapter 2

Matrices

Matrices are two-dimensional arrays that offer a concise way to store essential information for any type of application. In this chapter, we develop ways to manipulate matrices and do matrix algebra that prove useful for managing and analysing many different types of data-sets. We will also find links between matrices and vectors and their algebra: the rows or columns of a matrix can be viewed as vectors, and a vector can be viewed as a row or column matrix.

2.1 Operations with Matrices

Reading

Section 2.1 in Larson *et al*, has important fundamental definitions that make matrix algebra possible.

Note that AB need not equal BA. As a result, some of the algebraic rules we are accustomed to do not hold: for example, we will find that when we expand $(A+B)(A-B)$ according to the distributive rule in the next section, the terms $BA - AB$ that arise may not give O, so these factors may not yield the usual "difference of squares" result, $A^2 - B^2$.

Note also that matrix multiplication is defined so as to be useful for applications, especially for systems of linear equations. Be sure you have mastered being able to transform a linear system into a single matrix equation: see p 49/50 of Ed 4 (or p 52/3 of Ed 5).

Note that if the rows and columns of a matrix are viewed as row or column vectors, then the process of finding the ijth entry of a matrix product involves taking the equivalent of the dot (or scalar) product of the ith row of the first matrix with the jth column of the second.

Activity

Section 2.1 Exercises:
do Ed 4: odd numbers, 3, 7-11, 15, 19-31, 37, 50, 51, 52;
or Ed 5: odd numbers 3, 7-11, 15, 21-33, 39, 54, 57.
Attempt more if you have time.

2.2 Properties of Matrix Operations

Before we use matrix addition and multiplication to manipulate data, we need to check that we have enough properties to be able to work much like we do real numbers. We find that many of the algebraic rules still hold, *but not all*: a warning to be careful when using matrix algebra.

Reading

Study Section 2.2 in Larson *et al.* An easy section, but very important: be careful of the algebraic differences referred to above, for example those illustrated by Examples 4 and 5, and discussed in the notes between.

Note also that we can write the dot product of two vectors as a matrix multiplication:

if $\vec{u}$ and $\vec{v}$ are both row vectors of equal length, then $\vec{u} \cdot \vec{v}$ can be written as the matrix product $\vec{u}\vec{v}^T$. Explain why it is necessary to transpose $\vec{v}$ in order to write this matrix mltiplication.

Activity

Section 2.2 Exercises 3, 5, 9, 13, 17, 18, 19, 22, 23, 26, 27.
Also do Ed 4: 29, 36, 39, 53;
or else Ed 5: 29 (use technology), 31, 38, 41, 55.

2.3 The Inverse of a Matrix

All real numbers except 0 have multiplicative inverses: 2 has multiplicative inverse $1/2$, because $(1/2)(2) = 1$. We use this inverse as a multiplier on both sides, to clear the 2 in an equation like $2x = 5$.

For matrix equations, inverses are helpful in an analogous way.

The matrix equivalent of the number 1 is I. The inverse of matrix A will be that matrix which multiplies by A on either side, left or right, to give I.

But we find that not all matrices have inverses. For size reasons a matrix must be square to have an inverse: but even being square is no guarantee that the matrix will have an inverse!

Reading

Study Section 2.3 of Larson *et al.*

Note

- the definition of the inverse of a matrix;
- that when the inverse exists it is unique;
- a method of finding inverses: we can row-reduce $[A : I]$;
- the form of the inverse of a 2×2 matrix at the bottom of p 69 of Ed 4 (or p 73 of Ed 5). You may want it often enough to make it worth remembering. Verify it by multiplying it by the original to get I.

Notice that the inverse of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ cannot exist if $ad - bc = 0$. This number $ad - bc$ is called the *determinant* of the 2×2 matrix. We will define determinants fully later, so that they do indeed determine whether or not A has an inverse.

Activity

Master Section 2.3 Exercises, odds 1-19, 21, 23 with technology, and 27.
Also do Ed 4: 29, 35, 37, 42, 43, 45, 46, 49, 50;
or else Ed 5: 29 (with tech) 31, 37, 39, 44, 45, 47, 48, 51, 52.

2.4 Elementary Matrices

This section is *optional reading*. Elementary matrices can be used to prove some useful and important theorems.

2.5 Applications of Matrix Operations

There are many ways in which matrices can be used to facilitate the storing, processing, transmission, and analysis of data. This section introduces just four of these:

- stochastic matrices of probabilities, used for prediction;
- coding matrices, used for cryptography;
- input-output matrices, used in economics, for industrial output;
- least squares regression, which is covered in Algebra & Calculus II.

Once more you have a choice: study one of the first three applications listed above.

Reading

Look through Section 2.5 of Larson *et al*: study either *stochastic matrices* or *cryptography* or *Leontief input- output models*.

Activity

Do at least two exercises from Section 2.5, on the type of application you selected.

2.6 Summary: Matrices

A great deal can be obtained by adopting a little distance and a broader perspective in hindsight. Reviewing summaries may help.

Remember that the rows and columns of any matrix can be viewed as row or column vectors in as many dimensions as they have, and our process of finding the ijth entry of a matrix product involves taking the equivalent of *dot or scalar product* of the ith row of the first matrix with the jth column of the second.

> **Activity**
>
> Check the Chapter Summary in Larson *et al.* Grossman's summary on Matrices is also included in the coloured pages of Appendix A in the Study Guide.

> **Activity**
>
> You may want to attempt some of the Review Exercises. Also look over the types of questions asked in the Project 1.

Chapter 3

Determinants

We know how to clear factor 3 in the equation $3x = 5$. We simply multiply both sides by its multiplicative inverse or reciprocal, $1/3$. We also say we *divide by 3*.

We can clear any factor except 0 in that way. 0 is the only real number that does not have a multiplicative inverse.

Let us apply a similar argument to matrix algebra: suppose we have matrix equation $AX = B$. Some square matrices A have multiplicative inverses, A^{-1}, but many do not. This means that we cannot always get rid of a matrix factor A in a matrix equation. And that is why we do not define matrix division: *it simply can't always be done!*

To be able to distinguish between those matrices that have inverses and those that do not, we now define a number called the *determinant* of a matrix, so that

- when the determinant of the matrix is 0, the matrix has no inverse;
- when the determinant is not 0, the matrix is invertible (non-singular).

3.1 The Determinant of a Matrix

First we define the determinant of a 2×2 matrix. Then we extend the definition to matrices of any size, using a process of expanding by cofactors.

> **Reading**
>
> Study Larson Ch 3.1. Note the determinant of a 1×1 matrix (the entry itself) in Ed 4, top of p 113 (or Ed 5, top of p 121).
>
> Find the determinant of a 2×2 matrix in Ed 4, p 112 (or Ed 5, p 120).
>
> Study how these are extended to define the determinant of a square matrices of any size.

Note that you can expand by *any row or column*: the answer will be the same. Of course we choose the easiest one: the row or column with the most 0's.

In the next section we will see how we can simplify rows and columns and/or create more 0's, to make this task even easier.

Note that the *cofactor* of any entry has *two parts* to it:

- a plus or minus, following a chequerboard pattern, and
- a determinant: that of the matrix left after deleting the row and column containing that entry.

Two warnings on shortcuts:

- Do not use the alternative forward and backwards diagonal method (p 117 of Ed 4, or p 125 of Ed 5) on matrices bigger than 3×3. It does not work because it omits many terms.
- The shortcut for triangular matrices (p 117/8 of Ed 4, or p 126/7 of Ed 5) works for entries on the *main* diagonal. Be careful if you want to apply it to a triangular matrix where the entries lie on the *backwards* diagonal: matrices of order 6 or more may need a sign adjustment.

Activity

Master Section 3.1 Exercises 1, 3, 9, 11, 17, 21, 23, 31.

Also do Ed 4: 35, 39, 41, 43, 47, 48, 51;
or Ed 5: 37, 43, 45, 47, 53, 54, 57, and 39 with technology.

3.2 Evaluation of a Determinant using Elementary Operations

Determinants of matrices bigger than 3×3 are often tedious to expand. We want shortcuts, and we know some features of a matrix make the determinant easier to find:

- a row or column with many 0's;
- upper or lower triangular form: the determinant is then simply the product of main diagonal entries.

We may want to simplify a matrix to produce features like those above. If we use row-operations to achieve that then we need to know what effect row operations have on the determinant.

Reading

Study Larson, Chapter 3.2, noting Theorem 3.3. Note that we may use *column* operations when simplifying a matrix to find its determinant. We could not do that when solving systems of linear equations because the rows, not the columns, represented the equations.

Activity

Section 3.2 Exercises, master odd numbers in Ed 4: 1-21, 29, 33, 37, 39 (or Ed 5: 1-21, 31, 35, 39, 41, 43). Try others too.

3.3 Properties of Determinants

Reading

Note the properties of determinants listed in each of the Theorems in Larson 3.3, particularly Theorem 3.7.
Note Example 2: when finding determinants, we draw out factors from one row or column at a time.
This means that if A has 3 rows, $10A$ has determinant $10.10.10 \det(A) = 1000 \det A$.

Activity

Master the Section 3.3 Exercises, 1-31, 32, 35.
Also do Ed 4: 37, 47, 48 (or Ed 5: 39, 49, 50).

3.4 Chapter Summary on Determinants.

You should browse through Section 3.4 (and Section 3.5 in Ed 5) to gain a quick overview of some ways in which determinants are useful.

Note, however, that you need not study Eigenvalues and other Applications of Determinants in this course: we cover those in Algebra & Calculus II.

Activity

You may find it useful to review the Chapter 3 Summary on pp 152/3. Grossman's summary on Determinants is also included in the coloured pages in the Study Guide Appendix A.

Note, however, that for this course you are not required to study eigenvalues, adjoints, and Cramer's Rule.

Activity

Try some of the Chapter 3 Review Exercises.
The Cumulative Test covers Chapters 1 to 3.

Chapter 4

Complex Numbers

We know that no real number squares to give a negative number. Hence there is no *real number* that is the square root of -1, or -4, etc.

In this chapter we *create* or *invent* new *numbers* whose squares are negative real numbers, and use them to build the set of *Complex Numbers*. We also note the forms of some commonly-used functions of complex numbers.

These so-called complex numbers provide a fascinating and powerful extension of our number system. We quickly discover ways of using their algebraic and geometric properties to solve equations, rotate figures in the plane, and describe some physical phenomena. Mathematicians, engineers and scientists have found complex numbers useful to describe and analyse a host of physical phenomena: electric circuits, for example.

In this course we will use complex numbers to solve some polynomial equations. In applications to real-world phenomena, if we can find complex-valued solutions to an equation, we can extract the real-valued solutions, where appropriate. In Algebra & Calculus II, we show how complex numbers can be used to solve linear differential equations: their complex-valued solutions are quite easy to find, and these are used to obtain the real-valued solutions.

4.1 Complex Numbers: Definition

Reading

Find Appendix B at the back of the Study Guide. It is an extract from the 3rd edition of Larson & Edwards, *Elementary Linear Algebra*. Your prescribed 4th edition of this text does not include this chapter.
Read Appendix B, Chapter 8.1, pp 431-438. Notice that wherever possible, we manipulate complex numbers the way we do real numbers, simply replacing i^2 by -1 every time it arises.

Notice how the complex numbers provide yet another representation of a point in a plane. We already have cartesian coordinates (x, y) and position vectors $x\vec{i} + y\vec{j}$ to describe points in the plane R^2. Now we have $x + iy$ to describe points in the Complex plane too, and these numbers offer quite different algebraic strengths.

Activity

Do Appendix B, Section 8.1 Exercises: odds 1-43, 51 and 56.
Do more if you can manage, and think about question 58.

4.2 Conjugates and Division of Complex Numbers

The complex numbers $a + ib$ and $a - ib$ are very closely associated. We call them a complex conjugate pair. The fact that when you multiply them, the answer, $a^2 + b^2$, is pure real, can be used to make division by complex numbers very easy: we simply multiply numerator and denominator by the conjugate of the denominator to make the denominator pure real.

> **Reading**
>
> Study Appendix B, Chapter 8.2, p 439 - 44.

Note also the link with vectors. Some but not all of the geometry of vectors will extend to complex numbers. We talk about a complex number as having *modulus*. This is equivalent to a vector in R^2 having *length* or *magnitude*.

Recall, in particular, that when you subtract two position vectors, $\vec{a} - \vec{b}$, you get the vector that joins their endpoints. Hence $||\vec{a} - \vec{b}||$ measures the distance between their endpoints.

Complex numbers z and w also define two points in the complex plane, so $|z - w|$ measures the distance between them;
eg
$|(3 + 2i) - (1 + i)|$ measures the distance between points $(3, 2)$ and $(1, 1)$;
$|z - 2|$ measures the distance between points $z = a + ib$ and $2 = 2 + 0i$.

We can use these ideas to describe a circle: for example, the equation of the circle with centre $(2, 1)$ and radius 3 is:

$$\begin{aligned}(x-2)^2 + (y-1)^2 &= 9 \text{ in cartesian form,}\\ ||\vec{r} - (2,1)|| &= 3 \text{ in vector form}\\ |z - (2+i)| &= 3 \text{ in the complex plane}\end{aligned}$$

where $\vec{r}$ is the variable vector (x, y) from the origin and z the variable number $a + ib$ or point (a, b).

> **Activity**
>
> Do Appendix B, Section 8.2 Exercises, 1, 3, 4, 11, 12, odds 13-23, 27, 31, 33.
> The notes above, on the geometric similarity to vectors, may help you with 31 and 32.
> Think about how you would prove the results in 29 and 30.
> Do 34 if you wish.

4.3 Polar Form and de Moivre's Theorem

We now build the geometric strength of complex numbers, by giving them polar notation, just as we do points (x, y) in the cartesian plane.

Study Appendix B, Chapter 8.3, in three parts:

Reading

Firstly study Appendix B, p 445 to mid 447. Be very competent at transforming between $a + ib$ form and polar form $r(\cos\theta + i\sin\theta)$. Sketch the point in the correct quadrant first, to ensure correct choice of angle or signs.

Activity

Do Section 8.3 Exercises, 1-4, odds 5-13, 17, 19, 23.

Reading

Then study mid 447–449, carefully. You need the following identities:

$$\cos(\theta_1 + \theta_2) = \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2$$

$$\sin(\theta_1 + \theta_2) = \sin\theta_1 \cos\theta_2 + \cos\theta_1 \sin\theta_2$$

Be sure to establish that

- when we multiply two complex numbers, their angles *add*;
- when we divide them, their angles *subtract*;
- when we raise a complex number to a positive integer power, the angle *multiplies by that power*.

The last result is de Moivre's Theorem: obtained by simply applying the first result repeatedly.

Activity

Note that as a result of the conclusions of Theorem 8.4, some more tedious complex number calculations of multiplication, division and powers can be done more easily in *polar* form than $a + ib$ form. These are even better done in Euler's exponential form, which we introduce in the next section.
Do a few exercises from Section 8.3, 27-44, 36 and 38. Do more when you have met Euler form. Gain enough experience to be able to use your judgment to decide in advance which method is likely to be easier or quicker.

Reading

Study Appendix B, p 450–452 on nth roots of complex numbers. These may be done more intuitively in Euler form, so re-visit the process and exercises once you have Euler notation.

Note that the nth roots of z have

- the same modulus (length): the nth root of $|z|$;
- arguments (angles) that differ by $\frac{2\pi}{n}$.

So exactly n of them will fit into one revolution of 2π radians, before they start to repeat.

To plot them it helps start with the easiest one (found by setting $k = 0$). It has argument $\frac{\theta}{n}$ where θ is $\arg z$. Be careful not to plot it at the same angle as z; nor at the same distance.

Conclude that there are exactly n different n^{th} roots: 3 cube roots, 5 fifth roots, etc. When plotted, they lie equally spaced by angle $\frac{2\pi}{n}$ around a circle of radius $|z|^{\frac{1}{n}}$.

Activity

Do a few exercises from Appendix B, Section 8.3: 27, 33, 35, 37. These are often more easily done in Euler form, so come back to them in the next section.
Try 57, 59 and 61, using the fact that the solutions x can be expressed as nth roots of some complex number.

4.4 Euler's Form of a Complex Number

Your text does not cover the material of this section, or the next two: these notes are your readings.

We want simple effective notation to shorten polar form.

> **Reading**
>
> Refer to Appendix B, p 448, Theorem 8.4.

Notice that the *angles* θ_1 and θ_2 behave like *exponents*:

- when we multiply z_1 and z_2, their angles add;
- when we divide, they subtract;
- when we raise z to a power, its angle is multiplied by that power.

To remind us this is how the angles behave, when doing complex number algebra, it seems it might be helpful to *write the angle as though it were an exponent*: only the angle θ though, not the modulus r, which behaves quite ordinarily.

Remind yourself by plotting z and showing a, b, r and θ on your diagram, why it is that if $z = a + ib$ has polar form

$$z = r(\cos\theta + i\sin\theta).$$

Suppose we write just the angle part $(\cos\theta + i\sin\theta)$ in exponential notation instead, as $e^{i\theta}$: ie

$$\cos\theta + i\sin\theta = e^{i\theta}$$

Then the polar form is simply $re^{i\theta}$, so we write

$$z = re^{i\theta}.$$

We refer to this as the Euler form of the number, or Euler notation.

So Theorems 8.4 and 8.5 on p 448/9 in Appendix B, become, in Euler notation:

Given $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$

$$\begin{aligned} z_1 z_2 &= r_1 r_2 e^{i(\theta_1+\theta_2)} \\ \frac{z_1}{z_2} &= \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)} \\ \left(re^{i\theta}\right)^n &= r^n e^{in\theta} \end{aligned}$$

This notation makes multiplication, division and powers easy:

Example 4.1 *Simplify* $(1-i)^5(\sqrt{3}+i)^3$

In Euler form, this becomes

$$\begin{aligned} (\sqrt{2}e^{-i\frac{\pi}{4}})^5\,(2e^{i\frac{\pi}{6}})^3 &= (\sqrt{2}^5 e^{i\frac{-5\pi}{4}})\,(2^3 e^{i\frac{3\pi}{6}}) \\ &= 32\sqrt{2}\,e^{\frac{-3i\pi}{4}} \\ &= 32\sqrt{2}\left(\cos\frac{-3\pi}{4} + i\sin\frac{-3\pi}{4}\right) \\ &= 32\sqrt{2}\left(\frac{-1}{\sqrt{2}} + i\frac{-1}{\sqrt{2}}\right) \\ &= -32-32i. \end{aligned}$$

which probably offers a neater method than having to multiply 5 brackets of $(1-i)$ *by 3 of* $(\sqrt{3}+i)$.

Example 4.2 *Express* $(\frac{1+i}{\sqrt{3}+i})^{12}$ *in the form* $a+ib$.

This can be done far more concisely in Euler form:

Convert:

$$1+i = \sqrt{2}e^{i\pi/4}$$

$$\sqrt{3}+i = 2e^{i\pi/6}$$

So

$$\left(\frac{1+i}{\sqrt{3}+i}\right)^{12} = \left(\frac{\sqrt{2}e^{i\pi/4}}{2e^{i\pi/6}}\right)^{12}$$

$$
\begin{aligned}
&= \left(\frac{1}{\sqrt{2}}\right)^{12} \frac{e^{i3\pi}}{e^{i2\pi}} \\
&= \frac{1}{2^6} e^{i\pi} \\
&= \frac{1}{64}(\cos\pi + i\sin\pi) \\
&= \frac{1}{64}(-1+0i) = \frac{-1}{64}.
\end{aligned}
$$

Note that Euler form will clearly help with multiplication, division and powers, but that it is not useful for addition or subtraction, where form $a + ib$ is best.

> **Activity**
>
> Now do Appendix B, Exercises 8.3, 27, 33, 35, 37, using Euler form. Be careful to choose the correct quadrant angle.
> Also prove the following important result:
> if $z = re^{i\theta}$ then $\overline{z} = re^{-i\theta}$.

Note on Euler notation:

We write the argument as an exponent, because it behaves like one. But why do we include i in the exponent, and why use base e?

For a number of reasons: there are a number of areas in which complex numbers behave like exponentials with base e. Taylor series is one, but these are done in Algebra & Calculus II.

Let us look at an argument from calculus, instead:

we know that for real b, if we differentiate e^{bx} we get be^{bx}, ie we get the original function back again, but with coefficient b.

But if we differentiate $\cos x + i\sin x$ with respect to x, we get

$$-\sin x + i\cos x = i(\cos x + i\sin x)$$

ie the original back again, but with coefficient i.

So the i behaves like b in an expression of form e^{bx}.

So $\cos x + i\sin x$ behaves like e^{ix} when we differentiate it.

A Geometric View of Complex Number Multiplication: Rotations

Note that if we use complex number $re^{i\theta}$ as a multiplier on another complex number z, it changes both the modulus and argument of z:

- cause addition of angle θ to the argument of z, and
- multiply the modulus of z by r.

We conclude that multiplier $re^{i\theta}$ will cause *rotation* through angle θ and *lengthen or shorten* by factor r.

So complex numbers are useful, as matrices are, for rotating and resizing images in the plane.

Example 4.3 *Check this effect by multiplying the point $1+i$, say, by -1, i, and $2i$, and plotting the four complex numbers on the same diagram.*

1. *The multiplier $-1 = 1e^{i\pi}$ will rotate the point through π but keep the distance from origin.*
2. *Multiplier $i = 1e^{i\frac{\pi}{2}}$ will cause no change in distance, but will rotate points anticlockwise through 90 degrees.*
3. *Multiplier $2i = 2e^{i\pi/2}$ will stretch by a factor of 2 and rotate anticlockwise through a right angle.*

n'th roots can also be done in Euler form:

Example 4.4 *Find the nth roots of ae^{ib}.*

To catch all n of them, write the argument (angle) in its most general form, not just b, but $b + k2\pi$, where k is any integer:

$$ae^{ib} = ae^{i(b+k2\pi)}.$$

Suppose the nth roots are of form $re^{i\theta}$. Then we must have

$$(re^{i\theta})^n = ae^{i(b+k2\pi).}$$

So

$$r^n e^{in\theta} = ae^{i(b+k2\pi)}$$

Equating moduli and arguments both sides, because they represent the same complex number, we get

$$r^n = a \text{ and } n\theta = b + k2\pi$$

so solving for r and θ gives

$$r = a^{1/n} \text{ and } \theta = \frac{b + k2\pi}{n}$$

where k is any integer, but in fact the angles repeat after k consecutive values, so there are only k distinct (different) roots.

So the nth roots of $ae^{i(b+k2\pi)}$ are

$$a^{1/n} e^{i\frac{b+k2\pi}{n}}.$$

Note that it looks as though we "take nth roots" in the conventional real number way, but on the generalised angle!

Example 4.5 *Solve $z^3 = 4\sqrt{3} + 4i$.*

The z-values are the cube roots of the right hand side, so we plot $4\sqrt{3} + 4i$ or $4(\sqrt{3} + 1i)$ and find its Euler form, then we generalise the angle:

$$\begin{aligned} 4\sqrt{3} + 4i &= 8\, e^{i\frac{\pi}{6}} = 8\, e^{i(\frac{\pi}{6} + 2k\pi)}. \\ \textit{Then } z^3 &= 8\, e^{i(\frac{\pi}{6} + 2k\pi)} \\ \textit{so } z &= 8^{1/3}\, e^{i\frac{\pi/6 + 2k\pi}{3}} \\ &= 2\, e^{i(\frac{\pi}{18} + 2k\pi/3)}, \ \ k \textit{ any integer.} \end{aligned}$$

These cube roots are all distance 2 from the origin, and have angles that start at $\pi/18$, and change by $2\pi/3$, as k changes.
Three consecutive values of k, say $0, 1, 2$, will give three different points, but thereafter the angles terminate at earlier points. So there are 3 cube roots equally spaced around a circle of radius the nth root of the original. We can plot them starting at angle $\pi/18$.

Of course if we can find one easy root, we can find the others geometrically:

Example 4.6 *Find the cube roots of -8.*

We know there is one cube root, the real-valued one, -2. Plot this one on the negative x-axis: then the other 2 complex cube roots follow geometrically, spaced evenly around the same circle.

(Alternatively, write -8 in Euler form as $8e^{i\pi}$*, generalise the angle, and then take power 1/3, etc.)*

> **Activity**
>
> Plot the three cube roots of -8 found above, deduce their moduli and arguments from your diagram, express each in Euler's form, $re^{i\theta}$, convert each one to $a+b$ form, cube them and check that you do indeed get -8.

If we cannot find an easy root to get started, we simply work algebraically.

> **Activity**
>
> Find the 4th roots of i. It is not easy to guess what they will be: there is no obvious 4th root. We can work algebraically rather than geometrically: plot i to see that its modulus is 1 and its argument is $\pi/2$. In Euler form, $i = 1e^{i\pi/2}$. Generalise the angle by adding $k2\pi$, then proceed to take 4th roots. When you have found 4 answers, check them by raising them to power 4.

> **Activity**
>
> Go back to Appendix B, Section 8.3 Exercises, and do 45, 47, 49, odds 55 - 61 and 64, using Euler notation wherever it is helpful.

4.5 Solving Equations

We now have a bigger number system than the reals R. The complex numbers C contain the real numbers, in that every real number can be viewed as a complex number with imaginary part 0: eg $2 = 2 + 0i$.

This proves very useful: sometimes the real roots of an equation are hard to find, if they exist at all. If that equation can be solved more easily in C because C has additional algebraic properties, then we can solve for the complex roots instead, and then extract the ones we want, eg those that are pure real.

Note that we can now solve *any* quadratic equation with real coefficients even when the discriminant $b^2 - 4ac < 0$.

We do this by*completing the square,* or else using the familiar quadratic formula with $\pm i$ for the square roots of -1:

Example 4.7 *Solve $z^2 + z + 1 = 0$.*

Using the quadratic formula,

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

we find

$$z = \frac{-1 \pm \sqrt{1-4}}{2} = \frac{-1 \pm \sqrt{-3}}{2}$$

We now view $\sqrt{-3}$ as $\sqrt{3}\sqrt{-1}$ and use the fact that the square roots of -1 are $\pm i$. We have two complex solutions:

$$z = \frac{-1 \pm i\sqrt{3}}{2} = \frac{-1}{2} \pm \frac{i\sqrt{3}}{2}.$$

Notice that the two solutions are a conjugate pair.

Alternatively*, we could complete the square, then factorise the difference of squares that results:*

$$\begin{aligned} z^2 + z + 1 &= z^2 + z + \frac{1}{4} + \frac{3}{4} \\ &= \left(z + \frac{1}{2}\right)^2 - \frac{3}{4}i^2 \\ &= \left(z + \frac{1}{2} - \frac{\sqrt{3}}{2}i\right)\left(z + \frac{1}{2} + \frac{\sqrt{3}}{2}i\right) \end{aligned}$$

which gives the same solutions as before.

Example 4.8 *Solve* $z^3 + z^2 = -z$ *for complex numbers* z.

Bringing all the terms to one side gives $z^3 + z^2 + z = 0$.

Since we have no formula to solve this, we want to fatorise the LHS, if possible. We notice common factor z *throughout and take it out to get:*

$$z(z^2 + z + 1) = 0.$$

Equating each factor to 0 gives $z = 0$ *or* $z^2+z+1 = 0$*, which we have solved above. The solutions are therefore* $z = 0$ *and the two solutions listed in the example above.*

Activity

How do you know that the equation $z^2 + 2z + 2 = 0$ has no solutions in R?
Solve it in C, ie find all complex solutions, by completing the square or using the quadratic formula. You should find solutions $-1 \pm i$. Notice that they are a complex conjugate pair.

Now let us look at other new techniques we can use with complex numbers:

In *R* there is no factorisation of a *sum* of squares, $a^2 + b^2$. Recall the *difference* of squares factorisation, however:

$$a^2 - b^2 = (a + b)(a - b).$$

Now note that in C, because $i^2 = -1$, we can rewrite $+1$ as $-i^2$. Then

$$a^2 + b^2 = a^2 - i^2b^2$$

making it a *difference* of squares which we can factorise as $(a + ib)(a - ib)$.

So we have a factorisation for a sum of squares in *C*:

$$a^2 + b^2 = (a + bi)(a - bi)$$

Being able to factorise a sum of squares is helpful for solving equations in C.

Example 4.9 *Solve* $z^2 + 4 = 0$ *in* C*, ie for all complex solutions* z.

This is a sum *of squares, and it can be made into a* difference *of squares by replacing* $+1$ *by* $-i^2$*:*

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

Equating each factor to 0 gives two soluions: $z = 2i$ *and* $z = -2i$*.*

Note that the complex solutions form a complex conjugate pair.

Alternatively *we can solve* $z^2 = -4$ *by replacing* -4 *by* $4i^2$*. Then, keeping in mind that both* $\pm i$ *square to give -1, we can conclude that* $z^2 = 4i^2$ *gives* $z = \pm 2i$ *as before.*

Or we can solve $z^2 = -4$ *by finding the two square roots of -4 (using its Euler form,* $4e^{i\pi}$ *and the methods of the previous section).*

Example 4.10 *Solve in* C*:* $z^2 + 2 = 0$*.*

This is of the same form as the example above and can be solved in the same ways.

For example, writing the LHS sum of squares as a difference of squares instead:

$$z^2 + 2 = z^2 - 2i^2 = (z + i\sqrt{2})(z - i\sqrt{2})$$

Setting this equal to 0 gives solutions $z = i\sqrt{2}$ *and* $-i\sqrt{2}$*, again a complex conjugate pair!*

A note on solving polynomial equations of higher degree:

We often want to make one side 0 and *factorise* the polynomial on the other side, because factors yield solutions.

If no factor is obvious, we may need trial and error, to get started. We can test easy integer values in the polynomial, looking for one that makes it 0. If we find one, say a, then we know that $(x - a)$ is a factor of that polynomial.

The values of a we usually try are those that are factors of the *constant* in the polynomial: eg for $z^4 + z^3 - 5z^2 + z - 6$, we would test the numbers $\pm 1, \pm 2, \pm 3, \pm 6$, hoping to get a 0 on substitution. Check that $z = 2$ and $z = -3$ yield 0, and conclude that two factors are $(z - 2)$ and $(z + 3)$.

We may or may not be able to find other factors in the same way; when we find one or more, we can then divide the polynomial by them to see what must still be factorised, hopefully more easily. Sometimes we can avoid long division in this step: the quotient can sometimes be obtained by careful inspection and/or a little trial and error.

Activity

Factorise $z^3+3z^2+9z-13$ using the method suggested above: show that $z = 1$ makes this expression 0, and hence $(z - 1)$ is a factor. Then use trial and error or long division to show that the other factor is $z^2 - 4z + 13$. Hence show that the four complex solutions of the equation

$$z^3 + 3z^2 + 9z - 13 = 0$$

are $z = 1$ and $-2 \pm 3i$. (Notice the complex conjugate pair again.)

Example 4.11 *Solve* $z^4 + 4 = 0$.

We could write this as $z^4 = -4$, *and notice that* z *is therefore a fourth root of -4, and so find the 4th roots using the methods of the last section.*

But there are other interesting methods:

We could factorise the sum of squares,

$$(z^2 - 2i)(z^2 + 2i) = 0$$

then solve $z^2 = 2i$ *and* $z^2 = -2i$ *for* z *by finding the two square roots of* $2i$ *and the two square roots of* $-2i$.

Confirm that these give the same answers as the 4th roots above.

Alternatively, *we can complete the square unconventionally, by adding a middle term. Of course we must subtract it again too.*

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

Now this gives a difference of squares, so it can be factorised:

$$\begin{aligned} z^4 + 4 &= (z^2 + 2)^2 - (2z)^2 \\ &= [(z^2 + 2) - 2z][(z^2 + 2) + 2z] \\ &= (z^2 - 2z + 2)(z^2 + 2z + 2) \end{aligned}$$

Completing the square in each bracket conventionally with the middle term, gives

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

which is fully factorised into linear factors at last. So the solutions are

$$z = 1 + i, 1 - i, -1 + i, -1 - i.$$

Note that the complex number solutions to polynomial equations with real coefficients always occur in complex conjugate pairs. Trace the processes in these examples that cause this to happen! Of course there may be other roots that are pure real, and stand alone - the complex conjugate of a real number is itself!

Also note that the **Fundamental Theorem of Algebra**, which we cannot prove here, assures us that every polynomial with real coefficients is factorisable down to linear and/or quadratic factors with real coefficients (though it does not tell us how).

Now quadratic factors can always be factorised into two linear factors in C, because we allow use of i to get around the problem of square roots of negative numbers: see Example 4.7 above. And from each linear factor, we can obtain a solution to the polynomial being 0.

Hence

- any polynomial with real coefficients is indeed factorisable into n linear factors in C;
- nth degree polynomials equations hence have n *complex* roots. Among those are any real roots, and any repeated roots.

> **Activity**
>
> Solve $z^5 + 4z = 0$ in C expecting 5 solutions. You may make use of the results above.
> Also solve $z^4 - 4 = 0$, by factorisation. Check your answers by finding the 4th roots of 4.

> **Activity**
>
> Lastly, build a polynomial equation of degree 6 that has solutions $1, 2, 2+i$ and $1-i$. You may leave it factorised. The solution is given below, but try it yourself first.

We can reconstruct the polynomial from its factors.

Solutions 1 and 2 must have come from factors $(z-1)$ and $(z-2)$, but we are given only 4 roots to a 6th degree polynomial equation, so two roots are missing. Pure real solutions 2 and -1 do not have conjugates which introduce further solutions, but the other two do. We can find the conjugates of the two complex roots, then express each as a factor instead, then multiply all 6 of the factors to get the full polynomial:

$$(z-1)(z-2)(z-2-i)(z-2+i)(z-1+i)(z-1-i)=0.$$

Since any constant multiple is a polynomial of degree 6 with those solutions, we could offer the following more general form, with k any non-zero constant:

$$k(z-1)(z-2)(z-2-i)(z-2+i)(z-1+i)(z-1-i)=0.$$

4.6 Functions of Complex Variable z

Sometimes we want to use functions of complex numbers, for example trigonometric, exponential and log, and powers, to describe and analyse the behaviour of real phenomena, or to solve equations.

In this section, we show briefly how familiar definitions can be generalised to define useful complex-valued functions of complex variables z.

The exponential function, e^z:

We have already defined what we mean by e^{ib} for b real. To define what we mean by e^z, ie e^{a+ib}, we simply extend that:

for $z = a + ib$, we want

$$e^z = e^{a+ib} = e^a e^{ib} = e^a(\cos b + i \sin b),$$

and so we simply define e^z as that.

Example 4.12

$$\begin{aligned} e^{i(-3-2i)} &= e^{(2-i3)} = e^2\, e^{-i3} \\ &= e^2\,(\cos(-3) + i\sin(-3)) \\ &= e^2\,(\cos 3 - i \sin 3). \end{aligned}$$

We can get an approximate value for this, in $a + ib$ form, by using a calculator.

The trig functions:

For real a, we defined $e^{ia} = \cos a + i \sin a$.
From this and the fact that $\cos(-a) = \cos a$ and $\sin(-a) = -\sin a$, we also have

$$e^{-ia} = \cos(-a) + i\sin(-a) = \cos a - i \sin a.$$

Writing these two facts together, we have

$$\begin{aligned} e^{ia} &= \cos a + i \sin a, \\ e^{-ia} &= \cos a - i \sin a. \end{aligned}$$

Adding and subtracting these gives:

$$\begin{aligned}\cos a &= \frac{e^{ia}+e^{-ia}}{2},\\ \sin a &= \frac{e^{ia}-e^{-ia}}{2i}.\end{aligned}$$

We know from the above that this holds for any *real* number a. But we now extend these formulae to *complex* numbers too, by simply defining for any complex number z:

$$\begin{aligned}\cos z &= \frac{e^{iz}+e^{-iz}}{2}\\ \sin z &= \frac{e^{iz}-e^{-iz}}{2i}.\end{aligned}$$

We know that the right hand side has meaning because of the definitions we have for e^{iz} and e^{-iz} above.

We then extend to definitions of the other trig functions of z too, in the natural way: $\tan z$ is defined as $\frac{\sin z}{\cos z}$, and the other 3 trig functions are defined as the reciprocals of these.

Example 4.13

$$\begin{aligned}\cos(3+4i) &= (1/2)(e^{i(3+4i)}+e^{-i(3+4i)})\\ &= (1/2)(e^{3i-4}+e^{-3i+4})\\ &= (1/2)(e^{-4}e^{3i}+e^{4}e^{-3i})\\ &= (1/2)(e^{-4}(\cos 3+i\sin 3)+e^{4}(\cos 3-i\sin 3))\\ &= (1/2)[\cos 3(e^{-4}+e^{4})+i\sin 3(e^{-4}+e^{4})]\ \ etc.\end{aligned}$$

The log function:

To define $\log z$ we write z in Euler form and generalise the angle, just as we did for nth roots:

Suppose $z = re^{it} = re^{i(t+2k\pi)}$, k any integer.

We want $\log z = \log\left(re^{i(t+2k\pi)}\right)$.

If we can use the $\log AB = \log A + \log B$, to separate the logs, and make use of the fact that $\log e^a = a$, it seems reasonable to expect that

$$\log(z) = \log r + \log e^{i(t+2k\pi)} = \log r + i(t+2k\pi).$$

We can evaluate this RHS, so we accept it as the definition of $\log z$. It can then also be shown that indeed the usual log laws hold.

Because k can be any integer, $\log z$ has infinitely many values when z is complex. And they never repeat, as we see in the example find below.

The *principal value* of $\log z$ is that particular one that has imaginary part y satisfying $-\pi \leq y \leq \pi$ (like the principal argument of a complex number).

Example 4.14

$$\begin{aligned}\log(-16) &= \log 16e^{i\pi} \\ &= \log 16e^{i(\pi+2k\pi)} \ \textit{when we generalise the angle} \\ &= \log 16 + i(\pi + 2k\pi) = \log 16 + i(2k+1)\pi\end{aligned}$$

So the values of $\log(-16)$ *are roughly* $2.77 + i(2k+1)\pi$*.*

The principal value of $2.77 + i\pi$ *occurs when* $k = 0$*.*

> **Activity**
>
> Plot the values of $\log(-16)$ in the complex plane.
> Show that they all have the same real part, ie lie on the same vertical line, but that there is a vertical step of 2π or roughly 6.28 between their imaginary parts.
> Mark the principal value of $\log(-16)$ on your graph. Why will the principal value always be the one nearest the x-axis?

Complex Powers

We know what we mean by the power z^n when the base z is complex, and the exponent n is any integer. We have now also defined the exponential function e^z when the exponent is complex but the base is the real number e.

We now want to define complex powers z^w more generally, to cover the case where the base and exponent are both complex-valued, as in $(2+i)^{1-i}$ or i^i.

We make use of the definition above:

$$e^{a+ib} = e^a e^{ib} = e^a(\cos b + i\sin b).$$

To apply this definition to z^w, we must change the base from z to e.

Taking the natural log we know that $\log z^w = w \log z$. Rewriting this in exponential form, we get

$$z^w = e^{w \log z}.$$

If we now write $w \log z$ in the form $a + ib$, then we can use the definition of e^{a+ib} above to finish off, as in the example below.

Example 4.15 *Evaluate i^i.*

Notice that this has complex base and complex exponent.
We change the base to e using $i^i = e^{i \log i}$, then simplify the exponent, aiming at form $a + bi$.

Starting with $\log i$*, we have*

$$\begin{aligned} \log i &= \log\left(1e^{i(\pi/2+2k\pi)}\right) \\ &= \log 1 + i(\pi/2 + 2k\pi) \\ &= 0 + i(\pi/2 + 2k\pi). \end{aligned}$$

This means that $i \log i = i^2(\pi/2 + 2k\pi) = -\pi/2 - 2k\pi$.

So $i^i = e^{i\log i} = e^{-\pi/2-2k\pi}$, k any integer.

Note that in this last step, we could write the exponent as $2k\pi$ instead of $-2k\pi$, because k represents any *integer, positive or negative: either will do.*

Substituting values for k gives different values for i^i:
$k = 0$ *gives 0.208 and* $k = -1$ *gives 0.000388, approximately.*
Of course there are many more values:
all real in this case, because no i terms remain in the final form.

Example 4.16 *Evaluate* $(1 + 3i)^{2-i}$.

This also has form z^w, z and w complex, but requires more effort to evaluate than i^i does, because the imaginary part of the exponent does not disappear.

Only study this if you are confident with all that has gone before.

Changing the base to e, we get $(1 + 3i)^{2-i} = e^{(2-i)\log(1+3i)}$.

Now to find $\log(1+3i)$, *we use the Euler form of* $1+3i$. *The modulus is* $\sqrt{10}$. *For simplicity write* t *for the argument, ie* $t = \arctan 3$.

Then $1+3i = \sqrt{10}\, e^{it} = \sqrt{10}\, e^{i(t+2k\pi)}$.

Hence $\log(1+3i) = \log(\sqrt{10}\, e^{i(t+2k\pi)}) = \log\sqrt{10} + i(t+2k\pi)$.

Multiplying this by $(2-i)$ *gives* $(2-i)[\log\sqrt{10} + i(t+2k\pi)]$.

Gathering real and imaginary parts gives

$$\log 10 + t + 2k\pi + i(2t + 4k\pi - (1/2)\log 10)$$

Putting this exponent back into the power gives

$$(1+3i)^{2-i} = e^{\log 10 + t + 2k\pi + i(2t+4k\pi-(1/2)\log 10)}$$

Separating real and imaginary parts, aiming at form re^{it} *gives*

$$e^{\log 10}\, e^t\, e^{2k\pi}\, e^{i(2t+4k\pi-(1/2)\log 10)}$$

$$= 10\, e^t\, e^{2k\pi}\, e^{i(2t-(1/2)\log 10)} e^{i4k\pi}$$

and the last factor is 1. (Check sine and cosine values.)

If we put in an approximate value of $t = \arctan 3$, *ie 1.25, then* $k = 0$ *gives an approximate principal value of* $7.75 + 34.00i$.

Other values follow by setting $k = 1, 2, 3$ ... *and* $k = -1, -2, -3$ ...

Activity

Try the above examples yourself: aim at minimal reference.

4.7 Summary: Complex Numbers

Try to see the similarities and differences between the topics you have covered. Think about the ways in which complex numbers behave like geometric vectors in the plane, and the ways in which they do not.

Activity

The Readings do not have a summary of all the aspects of Complex Numbers that we have covered.

Review the main points, and list them, to reinforce your learning.

Module II

CALCULUS

Overview

The development of calculus took over 2000 years and is arguably one of the finest achievements of the human mind. The subject had its origins in the work of Archimedes (c. 287-212 BC) and others, and was further developed into its more modern form by Newton, Leibniz, and others in the seventeenth century. Since then calculus has been refined and applied to almost every area of human endeavor. Anything that represents change can be modelled with calculus and so you will find it used in engineering, science, business and even in art and sport.

In this module we begin with a revision of important aspects of *functions* and then progress to the foundations of *differential* and *integral* calculus. We cover much of the material from Chapters 1 to 6 of the textbook by Stewart, but some sections are omitted or covered only a little. As we come to each section of this module, you will be guided further on how carefully to study each section of the textbook.

The approach used here is to explore topics and concepts using a range of geometrical, numerical and algebraic techniques. Consequently there will be much emphasis in the module on using the computer package MATLAB or the equivalent technology, as an aid to exploring the graphics and numerics associated with development and understanding of concepts.

The theme in this Module is to represent functions and models using four techniques: tables; graphs; formulas; and words. That is we use numeric, geometric, algebraic, and verbal methods wherever possible.

> **Reading**
>
> Background Reading:
> Read the Preview of your calculus textbook **A Preview of Calculus**.

Chapter 1

Functions and Models

INTRODUCTION

You will need to work rather quickly through this chapter. Most of the functions in this chapter will be known to you from high school, although the section on inverses may be new or unfamiliar. The approach taken to these functions, and the way they are modelled, will probably be new to you.

This chapter on functions is very important for your success in developing calculus. All of the so-called elementary functions (powers, polynomials, exponential, logarithmic, and trigonometric) to be used in the course are introduced here.

The important thing for you is to become familiar with graphical and numerical approach to functions which adds so much meaning to common algebraic methods. The computer package MATLAB or its equivalent, is an excellent aid to this. You need to be able to think more graphically about functions and to be able to read and interpret tables of numerical data. This will help you to apply functions in the modelling of real-world phenomena.

There is also a section on constructing new functions from old. How to shift, flip and stretch graphs.

1.1 Four Ways to Represent a Function

Four ways are used here to represent functions. Remember that this means that you represent functions by tables, graphs, formulas, and words.

Reading

Study the set textbook (Stewart) edition 3, chapter 1 section 1.1, pp 11-21.

Note the use of the terminology domain/range, independent/dependent variable, increasing/decreasing. Also, the interval notation $[a, b]$ and (a, b) are introduced early in the section. For example: The set of x values such that $1 \leq x < 3$ is written as $\{x|1 \leq x < 3\}$ or simply as $[1, 3)$. Using symmetry by recognising *even* and *odd* functions can sometimes simplify things.

Activity

Do Stewart exercises Sec 1.1 nos 1, 5, 7, 9, 15, 61, 63.
Note that answers are given in the back for many of the odd numbered questions

Computer Activity

Stewart: section 1.4, no 21.

MATLAB note: for example

```
x=0: 0.01: 5;
f=10*x.^2;
g=x.^3/10;
plot(x,f,x,g)}
```

Note: for all of these Computer Activities, you are urged to use MATLAB or Scilab. You may of course use any other mathematics package, but we will only support MATLAB and Scilab. Of course you could instead use a graphics calculator or even a scientific calculator to do these Computer Activities, but it will take longer and is more tedious.

1.2 Mathematical Models: a Catalog of Essential Functions

In applications, mathematical modelling using functions is the most common way of representing technical phenomena in the real world. This section introduces modelling and gives a quick look at the important functions to be used by us in this course.

Linear functions have the property that the function changes by equal amounts for equal changes of the independent variable.

Reading

Stewart: study section 1.2, pp 25-29.

Activity

Stewart: do sec 1.2 exercises 5,7,11,13.

Computer Activity

A linear supply curve is given by $p = 1.05q + 0.02$, where q is the quantity of goods and p is the price of the goods. The demand curve is $p = 5.4 - 1.1q$. Using a computer plot, find a close approximation to the equilibrium values of p and q where the supply and demand curves intersect. Matlab hint: define a q set of points between (say) 0 and 10 using `q=0:0.01:10;`. This defines a set from 0 to 10 in steps of 0.01. Now the two graphs can be plotted one at a time using `hold on` to hold the first plot. For example `plot(q,1.05*q+0.02)` , and when you have plotted the second graph you can use `zoom` and then click on the intersection a few times to read off an accurate intersection. Check that the intersection is at about $q = 2.5$ and $p = 2.65$.

Polynomials should be quite familiar to you. We normally make use of polynomials up to about power 4. The most commonly used polynomials are linear functions (degree 1) and quadratic functions (degree 2).

Reading

Stewart: study sec 1.2, pp 29-31.

Activity

Stewart: sec 1.2 exercises 3, 9.

Computer Activity

Stewart: sec 1.2 no 27.

Power Functions are simply a power of x. For example x^3 or $\sqrt{x} = x^{0.5}$ or $x^{-1} = 1/x$. The key point in this section is how different powers compare on small and large scales. Basic shapes of graphs of power functions are developed. Note that it will be seen later that exponential functions grow (and collapse) much more quickly than power functions.

Reading

Stewart: study sec 1.2 pp 31-32.

Activity

Stewart: do sec 1.2 exercises 1(a),(c),3.

Computer Activity

Stewart: do sec 1.4 ex 25.

Rational Functions and Algebraic Functions are now defined.

Reading

Stewart: study sec 1.2, pp 32-33.

Trigonometric Functions. This section should be a revision for you of high school. Make sure that your calculators are set on radians. When we write $\sin x$, you must appreciate that x is a number, and so must be in radians. You should not need to use degrees mode on your calculator when x is a number. You can only use degrees mode when x represents the number of degrees of rotation, which is not such a good idea in mathematics, since mathematics is based on the study of number. Make sure that you can sketch the sine curve, showing the amplitude and period. The cosine curve is then just the same curve shifted back a distance of $\pi/2$. The tan curve is more complicated, since it has asymptotes at intervals of π. The identity $\cos^2 t + \sin^2 t = 1$ is not mentioned here, but is crucial, since it enables you to swap from cos to sin since then $\cos t = \pm\sqrt{1 - \sin^2 t}$ and $\sin t = \pm\sqrt{1 - \cos^2 t}$. The inverse trig functions will be seen later.

> **Reading**
>
> Stewart: See Reference inside front cover and sec1.2 pp 33-34, and Appendix C on Inverse Trigonometric functions (see Stewart page A25-A27).

> **Activity**
>
> Use a calculator to find one of the solutions to the equation $2 = 5\sin(3x)$. (Answer: $\sin(3x) = 2/5 = 0.4$, and so $3x = \sin^{-1} 0.4 = 0.4115$ (from calculator). Hence, $x \approx 0.1372$. If you didn't get that, check if your calculator is set in radians. Note: in MATLAB the inverse sine function $\sin^{-1} x$ is written as asin(x).

> **Computer Activity**
>
> Stewart: sec 1.4 exercise 9.

Exponential, Logarithmic, and Transcendental Functions are briefly defined and introduced, to be seen in detail later.

> **Reading**
>
> Stewart: study sec 1.2 pp 34,35.

1.3 New Functions From Old Functions.

In this section a summary is given of translations, shifts and stretches, reflections, combinations and composite functions (function of a function). The composite function $(f \circ g)(x) = f(g(x))$ is introduced. This is quite often called a function of a function for obvious reasons. For example, if $f(x) = x^2$ and $g(x) = \sin x$, then $(f \circ g)(x) = f(g(x)) = g^2 = \sin^2 x$. Composition of functions is one of the most useful elements of the study of functions.

Reading

Stewart: study sec 1.3 pp 38-45.

Activity

Stewart: do sec 1.3 ex 3, 5, 27, 35, 50.

1.4 Graphing Calculators and Computers

If you are using a graphics calculator, this section should be of interest to you. If you are using MATLAB you would be better to read the Booklet on an Introduction to MATLAB, by Patricia Cretchley. In terms of controlling the window of the plot in MATLAB you can use the command `axis equal` in order to make the scales the same on the x and y axes. If you just want to show the plot for the domain $-2 \leq x \leq 3$ and for the range $-1 \leq y \leq 4$, the MATLAB command is `axis([-2 3 -1 4])`.

Reading

Stewart: peruse sec 1.4 pp 49-53. Example 9 on p53 illustrates how a difficult equation can be accurately solved by graphical means using the zoom facility. In MATLAB just click on the zoom button on the toolbar of the plot and then click on the portion of the plot which you want to enlarge.

1.5 Exponential Functions

Exponential functions have the form $f(x) = a^x$ where a is a positive constant. Note that x is in the exponent, whereas for *power* functions x^a, x is in the base. Exponential functions have a simple property. The value of the function changes by equal ratios over equal intervals. That is, as the independent variable changes by a fixed amount, the value of the function changes by a constant factor. For example, given $f(x) = a^x$, then $f(x+3) = a^{x+3} = a^x a^3 = a^3 f(x) = Cf(x)$, and $f(x+6) = a^{x+6} = a^{x+3}a^3 = a^3 f(x+3) = Cf(x+3)$.

Reading

Stewart: study section 1.5 pp 55-61

Take particular note of the general graph of the exponential function a^x exemplified by figure 2 for the case when $a = 2$. You should be able to recognise the important features, ie the curve has a horizontal asymptote for the negative x axis, it then passes through the y axis at $y = 1$ and then climbs rapidly. Take note of Figure 4 also. For example, if $a = 0.5$, then the graph is $y = \left(\frac{1}{2}\right)^x = 2^{-x}$.

Note that the x (or t) is the independent variable and often denotes *position* (or *time*) for the function a^x (or a^t). Take note of families of exponential functions shown in Figure 3, the laws of exponents, and applications. The number e is very special. Remember that it has a value of about 2.718, and that it is just the value of a which makes a^x have a slope of 1 at the point (0,1) on the y axis. We will see soon that the function $y = e^x$ is the only function whose slope at any point (x, y) on the curve, is equal to the value of the function itself. In other words, the rate of change of e^x is equal to e^x. It is the only function in the entire universe that has this property.

Activity

Stewart: do sec 1.5 exercises 13,17,19.

Computer Activity

Stewart: do sec 1.5 exercise 25.

1.6 Inverse Functions and Logarithms

This section may be somewhat new to you. The concept of an inverse function is important. Make sure that you understand how to use tables and graphs to represent inverses and take particular note of the way to get formulas for inverses.

Reading

Stewart: study sec 1.6 pp 63-67.

Activity

Stewart: do sec 1.6 exercises 3,15,19,21.

Note that the graph of an inverse function is the mirror image of the function itself reflected about the line $y = x$.
Another neat property of inverse functions is that they undo one another algebraically. Thus $f(f^{-1}(x)) = f^{-1}(f(x)) = x$. For example if $f(T) = 4T - 160$, then $T = \frac{f}{4} + 40$, ie the inverse function is $f^{-1}(T) = \frac{T}{4} + 40$.
Then $f(f^{-1}(T)) = 4f^{-1}(T) - 160 = 4(\frac{T}{4} + 40) - 160 = T$. Also $f^{-1}(f(T)) = \frac{f(T)}{4} + 40 = \frac{4T-160}{4} + 40 = T$.

Logarithms are inverse functions for exponential functions. Make sure that you can sketch the graph of a log function. The easiest way to do that is to remember how to sketch the exponential function, and then just take the mirror image about $y = x$.
Take particular note of the three main laws.
Note that power functions with positive exponents always increase more rapidly than log functions.
We mainly concentrate on the natural logarithm function to base e and will write it as $\ln x$. Note that MATLAB uses the notation `log(x)` as the natural logarithm function.
We won't normally need to change to other logarithm bases, but if you ever need to do that, remember to use formula 10 on the bottom of page 69 of Stewart.

Reading

Stewart: study sec 1.6 pp 67-71.

Activity

Stewart: sec 1.6 do exercises 23, 25, 37, 49.

Chapter Summary The last section of this Chapter, on parametric curves, is to be omitted. It will be studied in the course MAT2100.
This first chapter is largely a revision of the function work you studied in high school. But the material forms a very important basis for the tools needed to develop the ideas and the applications of calculus.

Chapter 2

Limits and Derivatives

INTRODUCTION

In this chapter, the derivative is introduced as an instantaneous rate of change. Initially, tangents and speed are used to motivate practical examples, and then the limit definition is developed. By the end of the chapter you should be able to:

(i) understand the notion of an average rate of change;

(ii) understand the notion of a limit and be able to find the limiting value for simple examples;

(iii) find derivatives numerically by taking arbitrarily fine difference quotients;

(iv) visualize derivatives graphically as the slope of the graph when you zoom in (using your computer);

(v) interpret the meaning of first and second derivatives;

(vi) understand the meaning of an antiderivative.

2.1 The Tangent and Velocity Problems

Most people have a good intuitive idea of how to represent a tangent to a curve at a point on the curve. The tangent line must only touch the curve at a single point. You can approximate the tangent line by taking another point on the curve and drawing the line segment between it and the given tangent point. The slope of this line segment is easy to calculate, and then you can just take the other point closer and closer to the tangent point to get better and better approximations to the exact slope of the tangent.

Instantaneous velocity can be interpreted in two fundamental and related ways. It is the limit of average velocities, or it is the slope on the position versus time graph.

Velocity is a measure of change. Change in position over time. If you were filming an accelerating car with a fixed background, you can tell where the car is on a single frame of the film. However, if you were asked to estimate the exact speed of the car on one of the frames, you would have to look at the next (or previous) frame to get some idea. Of course you would also need to know the time between frames. So it takes two frames to estimate the speed as distance divided by time . If the frames of the film are very close together in time, the estimate of speed would be more accurate. The exact speed would be some sort of limit of a hypothetical camera as the time between frames got shorter and shorter.

Reading

Stewart: study sec 2.1 pp 93-97.

Activity

Stewart: sec 2.1 do exercises 3, 5, 7.

2.2 The Limit of a Function

We look now at the limit of a function $f(x)$ as x approaches a particular value a. Now, you might expect that the answer is always simply $f(a)$. But this is not always the case. For a simple function, such as $f(x) = x^2$, the limit as x approaches 2 will of course be $2^2 = 4$. But if $f(x) = \frac{x^2-4}{x-2}$, there is a problem! In fact $f(2)$ does not even exist in this case. But even so, you will find that $\lim_{x\to 2} f(x) = \lim_{x\to 2}(x + 2) = 4$. To get this, I wrote $x^2 - 4 = (x - 2)(x + 2)$ and cancelled the $x - 2$.

In this section we use numerical tables and graphics to see what the value of limits might be.

Reading

Stewart: study pp 98-105.
Example 4 shows that there are some dangers in just guessing the value of a limit by using a table of some values near the limiting point. Figure 7 illustrates the complex behaviour of the function when you get near $x = 0$. Strictly, we require that the *left-hand* and *right-hand* limits are the same. Then we say that the limit of the function exists and is given by their common values. Examples are given of some weird and wonderful functions. **Note:** the precise definition of a limit is given in Appendix D. It is indeed precise, but is a rather abstract definition, and is not examinable.

Activity

Stewart: sec 2.2 do exercises 1, 3, 13.

Computer Activity

Stewart: sec 2.2 exercise 23. From your answer to part (a), look back at sec 1.5. For part (b), plot for values of x from -2 to 3.

MATLAB **Note** for sec 2.2 exercise 23 part (b)

```
h=0.001; x=-2:h:3; y=(1+x).^(1./x); plot(x,y)
```

2.3 Calculating Limits Using the Limit Laws

In this section, we develop methods for finding the exact value of limits. To do this we basically make use of five common sense Limit Laws, and the Direct Substitution Property for simpler functions.

Reading

Stewart: sec 2.3 study pp 108-113.
If we are lucky, the value of $\lim_{x \to a} f(x) = f(a)$. The Direct Substitution Property is used here. But there are many weird functions for which it is not so easy. Then we need some tricks. Some of these tricks are exposed in Examples 1 to 7. **Note:** don't worry too much about the last part of this section from the start of Example 9 to the end of the section. You might read it for interest, but it is not examinable.

Activity

Stewart: sec 2.3 do exercises 1(a,f,g,h), 9.

2.4 Continuity

Continuity of a function at a point can be thought of geometrically as no sudden and abrupt change of position near or at the point. The shape can change at the point, but there can be no breaks there. A good example of this is seen in the function $|x|$ shown in Stewart:Figure 3, section 2.3. The function is continuous at $x = 0$, even though there is a dramatic change in shape there. Numerically, we say that values of the function near or at the point change gradually.

Reading

Stewart: sec 2.4 study pp 117-122.
Take note of Definition 1 and work through Examples 1 and 2. Note Definition 2 and 3 and work through Example 4. Theorems 4 and 5 are intuitively to be expected and the proofs are easy. Study Example 5 and take note of Theorem 7 which you will expect to be true anyway from their graphs. The rest of the section may be omitted, but should be read for interest by students majoring in mathematics.

Activity

Stewart: sec 2.4 do exercises 1, 3, 15.

2.5 Limits Involving Infinity

The important thing in this section is that you understand vertical asymptotes where $\lim_{x\to a} f(x) = \infty$ and that you understand horizontal asymptotes where $\lim_{x\to\infty} f(x) = L$.

Reading

Stewart: sec 2.5 study pp 128-137.
Make sure you understand that ∞ is not a number, but is a symbol for the idea that the process is getting bigger without bound. Don't worry about the precise definitions given in the Appendix D of the textbook. The ten Examples give good illustration of the concepts. You might omit Example 6.

Activity

Stewart: sec 2.5 do exercises 3, 21(hint: see Example 5), 47(b).

Computer Activity

Stewart: sec 2.5 do exercise 11, 13.

2.6 Tangents, Velocities, and Other Rates of Change

For tangents, you probably learnt the value of the *slope* as "the rise over the run". If you are given a curve or line whose equation is $y = f(x)$ and you take a "run", or change in x, of $\Delta x = h$, then the "rise" is the change in y given by $\Delta y = f(x+h) - f(x)$. Then the slope of the tangent is approximately $\frac{\text{rise}}{\text{run}} = \frac{\Delta y}{\Delta x} = \frac{f(x+h)-f(x)}{h}$ and the approximation gets better and better as $\Delta x = h$ gets smaller.

For moving objects travelling at a constant velocity, the velocity is found by calculating $\frac{\text{displacement}}{\text{time}}$. The displacement, or distance from the starting point, is normally given as a function of time, ie $f(t)$. When the object's velocity is changing, but we wish to calculate its velocity at a particular time $t = a$, we have to get smarter. In that case we take a value of t near a, say $a + h$, and find the average velocity for t between a and $a + h$. This is just $\frac{f(a+h)-f(a)}{h}$, ie the distance travelled between $t = a$ and $t = a+h$. Then we shrink h and find the limit as $h \to 0$ giving the instantaneous velocity at $t = a$.

Reading

Stewart: sec 2.6 study pp 139-145.
Take particular note of the Δ notation introduced on p143. Δ is the Greek letter corresponding to our "d", and signifies difference, or change. **Note:** the ratio

$$\frac{\Delta y}{\Delta x} = \frac{f(x+h) - f(x)}{h}$$

is called a *difference quotient*. We will often use this term.

Activity

Stewart: sec 2.6 do exercises 3, 5(a,b), 15, 25.

Computer Activity

Stewart: sec 2.6 do exercise 5(c).

2.7 Derivatives

Rates of change occur in almost all phenomena in our lives and throughout our immediate environment and the rest of the universe. Given a function $y = f(x)$, we denote by $f'(a)$ the rate of change of f at a particular value of $x = a$. We call $f'(a)$ the *derivative* of The function f at a. By the way, the study of derivatives is called *Differential Calculus*. So differential calculus is just about how things change. In Chapter 5 we will start on *Integral Calculus*, which is another story.

Reading

Stewart: sec 2.7 study pp 148-152.
Work through Examples 1, 2,3,4. Take particular note of Examples 5 and 6. Here you can see how to use the *difference quotient* to help you interpret the units of the derivative which is an important aid to explaining the actual contextual meaning of the derivative in different practical situations.

Activity

Stewart: sec 2.7 do exercises 1, 7, 29, 31.

Computer Activity

Stewart: sec 2.7 do exercise 11.

2.8 The Derivative as a Function

We have seen above how to find the derivative $f'(a)$ at a particular point. If we can do that at every point in an interval, for example $-2 \leq x \leq 3$, we will then know the derivative function $f'(x)$ over that domain of x values.

Reading

Stewart: sec 2.8 study pp 155-165.
In Example 1, the slope is found at each point of Figure 1 and thus the derivative is known and the sketch of the derivative function follows as shown in Figure 2(b). Example 2 shows how to use difference quotients to calculate good approximations to the value of the derivative at a collection of points, and thus the derivative function can be sketched. Do Example 3, don't worry too much about Example 4, and Example 5 is a little harder. Take particular note of alternative notations for the derivative. The Leibniz notation $\frac{dy}{dx}$ is a particularly good one, since it reminds us that

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Pages 160-162 are a commentary on differentiability. Essentially, a function $f(x)$ is differentiable on some domain like $-2 \leq x \leq 3$, if the graphical representation is smooth. Thus the plot can have no abrupt change of character, since the rate of change is smooth. Take particular note of *higher derivatives* and their notations and interpretations.

Activity

Stewart: sec 2.8 do exercises 1, 19, 37, 39.

MATLAB Note: given the function $y = f(x) = e^{x^2}$ the following code will plot, between $x = 0$ and $x = 1$, values of the derivative function $f'(x)$ which just happens to be $f'(x) = 2xe^{x^2}$. First of all the function f is plotted, and then the function f' is plotted in red for comparison. Note how MATLAB does comments. Anything in the line after the % is taken as a comment. Note also the need to use `.*` and `.^` when multiplying or taking powers of an array of numbers, such as `x`.

```
h=0.01; % the step size for x
x=0:h:1; % the set of x values
y=exp(x.^2); % the set of y values
plot(x,y);
hold on; % holds the plot to overlay the next one
deriv=2*x.*exp(x.^2); % the derivative values
plot(x,deriv,'r') % plot the derivatives in red
```

In the above, we have used the known function for the derivative. If we don't know the derivative function and we want to find approximate values for the derivative, we can use the difference quotient $\frac{f(x+h)-f(x)}{h}$ at a set of x values all at one time using MATLAB . The trick is to calculate the values of $f(x)$ and then to calculate the values of $f(x + h)$ by shifting x to $x + h$. Then the difference quotients can all be calculated at the same time and plotted as follows:

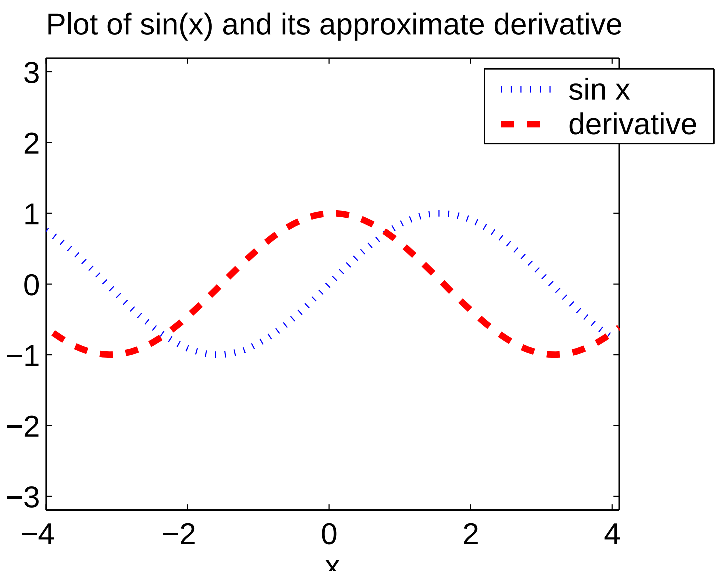
```
h=0.01; % the step size for x
x1=0:h:1; % the set of x values
y1=exp(x1.^2); % the set of y(x) values
plot(x1,y1);
hold on; % holds the plot to overlay the next one
x2=x1+h; % shifts all the values of x1
y2=exp(x2.^2); % the set of y(x+h) values
deriv=(y2-y1)/h; % the approx derivative values
plot(x1,deriv,'r') % plot the derivatives in red
```

Or, if an even shorter way of coding is used:

```
h=0.01; % the step size for x
x=0:h:1; % the set of x values
plot(x,exp(x.^2)); % plots the function
hold on;
% now plot approximation to the derivative in red
plot(x,(exp((x+h).^2)-exp(x.^2))/h,'r');
```

Computer Activity

Problem: given the function $y = f(x) = \sin x$ for $-4 \leq x \leq 4$, use technology to approximate the derivative using step size $\Delta x = h = 0.1$, and then plot the function in blue together with the derivative in red. (See figure following for the answer.)
Stewart: sec 2.8 do exercise 41.



2.9 What does f' say about f?

If we know something about the derivative function $f(x)$, how can we find out about the function itself $f(x)$? So we need to be able to reverse our understanding of the rate of change process. For example, if someone measured the speed of your car at certain times and you then needed to find out where the car was at those times, you would need to reverse the rate of change process. Formally we call this *antidifferentiation* and a knowledge of how to do this turns out to provide very powerful tools for solving problems. This section gives us a graphical introduction to this reversing process.

Reading

Stewart: sec 2.9 study pp 168-172.

Activity

Stewart: sec 2.9 do exercises 1, 7, 27.

Computer Activity

Stewart: sec 2.9 do exercise 29.

MATLAB **Note on exercise 29 above.**

```
x=linspace(0,4*pi,2000);  % 200 points between 0 and 4pi
deriv=exp(-x/10).*sin(x);
plot(x,deriv);
grid; % puts a grid onto the plot
xlabel('x'); % how to do an axis label
title('Now hand plot the function which gives this derivative');
```

Chapter Summary
By now you should have a good idea of the notion of a derivative from a graphical, numerical, and algebraic point of view. Limits are at the basis of this understanding. It is also important for you to be able to use the difference quotient to approximate the value of derivatives and also to interpret practical meanings of derivatives and to find the corresponding units of the derivative function.

Chapter 3

Differentiation Rules

INTRODUCTION

In this chapter some fundamental rules are developed for working with derivatives. If we had to calculate the limit every time we wanted the derivative of a function, life would be rather boring and inefficient. It was important for you to know that the derivative $y' = f'(x)$ is a limit of a difference quotient, for example:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x};$$

or the same thing

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

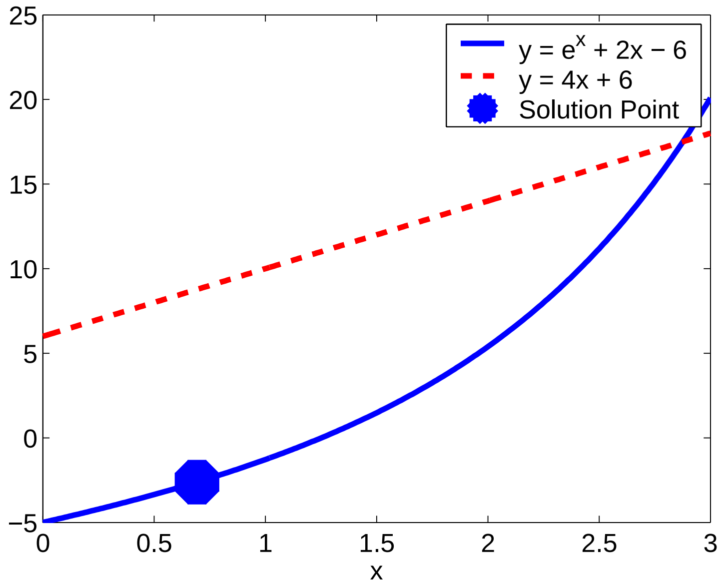
But now we develop some rules for making things easier. Firstly we determine the rules for finding the derivative of the elementary functions: polynomials, powers, exponential, trigonometric, and logarithmic. Along the way we develop rules for finding the derivative of sums and differences of functions, products and quotients of functions. The most important differentiation rule of all is how to find the derivative of a function of a function. For example, you were told in Section 2.8 that the derivative of $y = f(x) = e^{x^2}$ is $y' = f'(x) = 2xe^{x^2}$. To do this we use the *chain rule* which will be developed in this chapter. Section 3.8 of the textbook is to be omitted. It will be studied in the next course.

3.1 Derivatives of Polynomials and Exponential Functions

This section is an introduction to finding the derivative function for sums of simple power functions and exponential functions.

> **Reading**
>
> Stewart: study sec 3.1 pp 183-190.
> You will not be required to be able to prove that $\frac{d}{dx}x^n = nx^{n-1}$, except for special cases as shown. The general proof uses the Binomial theorem which should be of interest to mathematics majors.
> It is of interest to note that the derivative $\frac{d}{dx}$ is a *linear operator* which means that the **Constant Multiple Rule**, the **Sum Rule**, and the **Difference Rule** can be combined into one to give
>
> $$\frac{d}{dx}\left(C_1 f(x) \pm C_2 g(x)\right) = C_1 \frac{d}{dx} f(x) \pm C_2 \frac{d}{dx} g(x).$$
>
> We will see in sec 3.5 that the derivative of an exponential function of the form a^x is $(\ln a)a^x$. When $a = e$, we get the simplest of all differentiation rules $\frac{d}{dx}e^x = e^x$. On page 189 of the textbook, e is defined in a very formal way as a limit. You should be able to recall that $e \approx 2.718$.

Remember that if you have an equation like $e^{x+1} = 3$, to solve for x you just take natural logs of each side to get $\ln\left(e^{x+1}\right) = \ln 3$, and since log is inverse to exponential, you then get $x+1 = \ln 3$, $x = \ln 3 - 1 \approx 0.0986$ (from my MATLAB `log(3)-1`). Then you can solve problems like the following:
Example
Find where the tangent to the curve defined by $y = e^x + 2x - 6$ is parallel to the line $y = 4x + 6$.
Solution The slope at any point on the curve is $y' = e^x + 2$ and the slope of the line is 4. Hence
$e^x + 2 = 4$, $e^x = 2$, $x = \ln 2 \approx 0.6931$. At $x = \ln 2$, $y = e^{\ln 2} + 2(\ln 2) - 6 = 2 + 2\ln 2 - 6 \approx -2.6137$, and so the tangent, at the point $(x, y) \approx (0.6931, -2.6137)$ on the curve, has slope 4. See following figure.



Activity

Stewart: sec 3.1 do exercises 3, 7, 11, 13 (remember that π is a constant), 23, 41, 45, 49.

Computer Activity

Stewart: sec 3.1 exercise 33, 50.

3.2 The Product and Quotient Rules

Don't be fooled into using $\frac{d}{dx}(f(x)g(x)) = f'(x)g'(x)$. It is a very common error. As you can see in this section the rule is

$$\frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x).$$

So in fact you must *differentiate each function one at a time and add the results*. This rule extends to products of more than two functions as follows:

$$\frac{d}{dx}(u(x)v(x)w(x)) = u'(x)v(x)w(x) + u(x)v'(x)w(x) + u(x)v(x)w'(x),$$

and so on. The quotient rule is useful and should also be learned. Don't forget to divide by g^2 and not just g. I sometimes forget!

Reading

Stewart: sec 3.2 study pp 193-198.
A note on notation: In Example 3 on p195, $g'(4) = 3$ means that the derivative $g'(x)$ at $x = 4$ is 3.

Activity

Stewart: sec 3.2 do exercises 1, 3, 5, 13, 29, 31, 37(c), 40, 41.

Computer Activity

Stewart: sec 3.2 exercise 23.

3.3 Rates of Change in the Natural and Social Sciences

In this section, we will only study the subsection on physics. If you are studying a major in chemistry, biology, business, or another science, you should be interested to see the applications to those areas.

Reading

Stewart: sec 3.3 study pp 200 to the top of p204.
The key concept in Example 1 is that if you are given the position s of an object as a function of time t, then the velocity is a measure of change. It is the rate of change of position with respect to time. The average velocity over a time period Δt is $\frac{\Delta s}{\Delta t}$ and the limit of this ratio as $\Delta t \to 0$ is the instantaneous velocity $v(t) = \frac{ds}{dt}$. The acceleration is rate of change of velocity $a(t) = \frac{dv}{dt}$ which can also be written as $v'(t)$ or $s''(t)$, or in the Leibniz notation $\frac{d^2s}{dt^2}$. Examples 2 and 3 give important applications to density and electric current problems. We will certainly see more about these applications in the following course MAT2100.

Activity

Stewart: sec 3.3 do exercises 1, 7, 13, 15, 17.

3.4 Derivatives of Trigonometric Functions

The derivatives of $\sin x$ and $\cos x$ are easy to remember. Check carefully to see the relationship between the graph of the sine function and its derivative, the cosine function.

Reading

Stewart: sec 3.4 study pp 213-218.
You will not be expected to formally prove that

$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

but the result should not be a surprise, since you already know that the slope of the graphs of $y = x$ and $y = \sin x$ are both equal to 1 as they pass through $x = 0$. Consequently, as $x \to 0$ if the graphs look the same, the ratio of x to $\sin x$ must approach 1.
The formal proof that $\frac{d}{dx} \sin x = \cos x$ and $\frac{d}{dx} \cos x = -\sin x$ will not be required.
The derivatives of $\sin x$, $\cos x$, $\tan x$, $\csc x$, $\sec x$, and $\cot x$ should be memorised.

Activity

Stewart: sec 3.4 do exercises 5, 7, 9, 31, 37.

Computer Activity

Stewart: sec 3.4 do exercise 19.

3.5 The Chain Rule

The chain rule is the most powerful rule in all of differentiation. So it is essential that you master it. In real life, things are hardly ever simple, so you will need to be able to handle rates of change of composite functions like $y = F(x) = e^{(x^2)}$. This is a function of a function. The inner function is $g(x) = x^2$ and the outer function is $f(x) = e^x$, so that $F(x) = f(g(x)) = e^{g(x)} = e^{x^2}$. The chain rule then is

$$\frac{dF}{dx} = \frac{dF}{dg}\frac{dg}{dx} = (e^g)(2x) = 2xe^{x^2}.$$

Reading

Stewart: sec 3.5 study pp 220-225.
You should take note of the informal proof on page 221, but you are not required to learn the formal proof on page 227. Take particular note of the power rule for composite functions given by formula 4 and the proof that the derivative of the exponential function a^x is $a^x \ln a$ given by formula 5.
The subsection on tangents to parametric curves is to be omitted.

Activity

Stewart: sec 3.5 do exercises 1, 3, 5, 17, 19, 23, 41, 43, 47, 59, 79.

Computer Activity

Stewart: sec 3.5 do exercise 63.

3.6 Implicit Differentiation

The idea of implicit functions is useful when an equation cannot be solved explicitly for y in terms of x or vice versa. For example, the circle equation is $x^2 + y^2 = 4$, but you cannot get y as an explicit function of x. The best you can do is $y = \pm\sqrt{4 - x^2}$. If you had an equation like $x^2y^3 - x^5y^7 + x^2 + x = 4$ you would have no hope. This is an implicit relationship between x and y and it is quite easy to find the rate of change $\frac{dy}{dx}$ or $\frac{dx}{dy}$ for such functions - as we shall see. By the way take note in passing that

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

Reading

Stewart: sec 3.6 study pp 232-238.
Note how important the chain rule is in implicit differentiation. Example 1 on pp 233 and 234 is a good introductory illustration of this.
Make sure that you can prove the formulas for the derivatives of the inverse functions $\sin^{-1} x$ and $\tan^{-1} x$. These are useful in calculus.

Activity

Stewart: sec 3.6 do exercises 1, 3, 7, 31, 33, 51.

Computer Activity

Stewart: sec 3.6 do exercise 21

3.7 Derivatives of Logarithmic Functions

The most important result in this section is the somewhat surprising fact that $\frac{d}{dx}\ln x = \frac{1}{x}$. The easiest way to prove this is as follows:
write $y = \ln x$. Then $e^y = e^{\ln x} = x$, ie $e^y = x$. Now differentiate implicitly with respect to x to get $e^y \frac{dy}{dx} = 1$, (by the chain rule). Hence $\frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$, as required.

Reading

Stewart: sec 3.7 study pp 240-244.
Example 6 is interesting. The derivative of $\ln|x|$ is $\frac{1}{x}$ and so is positive when x is positive and negative when x is negative. This result will be used when we get to integration in the following chapters.
Take note of logarithmic differentiation and its use in finally proving the power rule, that $\frac{d}{dx}x^n = nx^{n-1}$ for all real numbers n.
Omit p255.

Here is one more example of logarithmic differentiation. Find the derivative of x^x, $x > 0$.

$$\begin{aligned}
y &= x^x \\
\ln y &= \ln x^x = x \ln x, \text{ then differentiate implicitly w.r.t. } x, \\
\frac{1}{y}\frac{dy}{dx} &= 1 \times \ln x + x\frac{1}{x} = \ln x + 1 \\
\frac{dy}{dx} &= y(\ln x + 1) = x^x(1 + \ln x)
\end{aligned}$$

Hence $\frac{d}{dx}x^x = x^x(1 + \ln x)$.

Activity

Stewart: sec 3.7 do exercises 3, 9, 17, 19, 25, 33.

Chapter Summary

Section 3.8 is to be omitted. Try some of the Review problems on pp 255-257.

In this chapter you have developed some important tools for differentiating functions. The most powerful of these is the chain rule. Familiarity with these rules is essential for later courses in solving differential equations and is immediately necessary in this course for the two chapters on integration.

Chapter 4

Applications of Differentiation

INTRODUCTION

In this chapter we make use of differentiation rules to develop some more powerful methods involving the chain rule and some tests and methods for solving optimisation problems. Some sections will either be omitted or covered very lightly. The last section on antiderivatives is an important lead into integration.

4.1 Related Rates

Related rates problems are practical applications of the chain rule. For example, if you want to know how fast the area of a circle is changing when the radius is 2 metres and if the radius is increasing at 0.5 metres per second, then we need to express our variables in composite form. In this case we write down the *knowns*: $A = \pi r^2$ and $\frac{dr}{dt} = 0.5$, and $r = 2$. In this case we have $A = f(r)$, where r is some function of t.

The chain rule gives:

$$\frac{dA}{dt} = \frac{dA}{dr}\frac{dr}{dt} = 2\pi r \times 0.5 = \pi r.$$

Now we substitute $r = 2$, to get $\frac{dA}{dt} = 2\pi \approx 6.82$ square metres per second.

Note that condition $r = 2$ cannot be used until we calculate some rates of change. Otherwise we would not have been able to find dA/dr, it would have been zero if r was set to 2.

Reading

Stewart: sec 4.1 study pp 263-266.
Work through each of Examples 1,2,3. The Strategy given on page 265 can be very helpful. Note the warning given about substituting known values before you have the chance to work the rates of change. Work through Examples 4 and 5.

Activity

Stewart: sec 4.1 do exercises 3, 5, 7, 11(note: mi means miles), 31*(note: * for harder problems), 37*.

4.2 Maximum and Minimum Values

Optimisation is about things like minimising costs and maximising output. Or minimising stress in a material and maximising safety. Or minimising distance while maximising delivery of goods. And so on.

Reading

Stewart: sec 4.2 study pp 269-274.
This section introduces some important terminology (ie language) for discussing maxima, minima, and other critical points. The Closed Interval Method on page 273 gives a good summary of the way to determine absolute extrema. Work through the seven Examples.

Activity

Stewart: sec 4.2 do exercises 5, 17, 19, 23, 27, 35, 37, 59*.

4.3 Derivatives and the Shapes of Curves

In this section we develop ideas on the shape of curves using the First Derivative Test and then use a knowledge of concavity to develop the Second Derivative Test for finding Local Maxima and Minima. The Mean Value Theorem is to be omitted, but should be read for interest by mathematics majors.

Reading

Stewart: sec 4.3 study pp 280-284.
Increasing/Decreasing functions, The First derivative Test, Concavity, and the Second Derivative Test, are all important. Make sure that you know when a curve is *concave up* (ie $f'' > 0$, or the slope is increasing), and when it is *concave down* (ie $f'' < 0$, or the slope is decreasing). Study Examples 2,3,4 and 5. You might just read Examples 6, 7, and 8 for interest.

Activity

Stewart: sec 4.3 do exercises 5, 7, 15, 21, 55*.

Computer Activity

Stewart: sec 4.3 do exercise 47*.

4.4 Graphing with Calculus and Calculators

In this section we will concentrate on the first example and officially omit the rest. However, there are some tips to be found if you care to read the rest.

Reading

Stewart: sec 4.4 study pp 289-290.
Do Example 1 only.

Computer Activity

Stewart: sec 4.4 do exercise 7.

4.5 Indeterminate Forms and l'Hospital's Rule

Omit this section.

4.6 Optimisation Problems

In this section, some principles are formulated for modelling problems and then using well constructed problem-solving techniques to find optimal solutions.

Reading

Stewart: sec 4.6 study pp 306-311.
First of all six important steps in solving optimising problems are formulated. You should work through each of Examples 1, 2, 3, 4, and 5.

Activity

Stewart: sec 4.6 do exercises 7, 9, 21, 39, 43.

4.7 Applications to Business and Economics

Omit this section. Unless you want to get rich quick. It will not be examined, but should be read with interest if you are doing a business major.

4.8 Newton's Method

Omit this section. It is included in a later mathematics course together with many other powerful numerical techniques.

4.9 Antiderivatives

If I tell you that $f'(x) = 2x$ and ask you to find the function $f(x)$, you should have no trouble. Because you remember the power rule for differentiation, you will tell me in a flash that $f(x) = x^2$. Careful! $f(x) = x^2 + 1$ also does the job. So does $f(x) = x^2 - 12345\pi$. The general answer is $f(x) = x^2 + C$ where C is an arbitrary constant.

In this section, we continue the process of thinking backwards from the known derivative. To do that we need a very good knowledge of the differentiation rules. In the above, I had to use the fact that $\frac{d}{dx}x^2 = 2x$ and that $\frac{d}{dx}C = 0$.

We need to think backwards when we know the rate of change of a process, but we want to find the actual state of the process. For example if you look at the speedo on a car, you can see what the velocity is, but you might want to calculate from it your displacement from home in a given time.

Finding antiderivatives means that you have started integrating.

Reading

Stewart: sec 4.9 study pp 327-332.

Activity

Stewart: sec 4.9 do exercises 1, 9, 15, 23(including finding the two constants), 27(careful when you get to x^{-1}), 39, 55*.

Computer Activity

Stewart: sec 4.9 do exercise 32.

Chapter Summary

Work through some of the Review problems, see pp 335-341.

In this chapter we have worked through a wide range of applications of the differentiation rules which were developed in the previous chapter. So you can already see that differential calculus can be very useful. But the story continues and the true power of calculus awaits us in the next two chapters where we introduce Integration.

Chapter 5

Integrals

INTRODUCTION

In a simple sense, differentiation is about rates of change. The difference quotient $\Delta y/\Delta x$ reminds us that the derivative is about ratios of quantities. In a sense, integration is even simpler. It is about summing things to find quantities.

In this chapter we develop the idea of a Definite Integral which is used to sum quantities. We find that the definite integral has a rather amazing connection with the derivative. The connection is called *the Fundamental Theorem of Calculus.* The rest of the chapter is devoted to the development of standard ways to find exact or approximate values of definite integrals.

5.1 Areas and Distances

This section is an important introduction to integration. The idea of calculating areas under curves using thinner and thinner rectangle approximations will lead to a more general way of summing a wide range of quantities (eg velocity functions).

Reading

Stewart: sec 5.1 study pp 343-352.
Note that the area can be approximated by rectangles using the left endpoint of each interval, or using the right interval. When you take the limit as $n \to \infty$, (ie $\Delta x \to 0$), you get the same answer either way. So it doesn't matter whether you use left or right endpoints, in fact you can use any sample of points, one in each subinterval, as discussed on page 348. To keep things simpler, we will only use left or right endpoints, or sometimes the midpoint. The last subsection on The Distance Problem shows that we can calculate approximations to the distance travelled in the same way, ie distance = velocity $\times$ time. If we assume that $v(t)$ is constant in each subinterval Δt, we get distance $= v(t) \times \Delta t$. Then we just sum all of these values across the domain and then take the limit as $\Delta t \to 0$, ie the number of subintervals $n \to \infty$. This will give us the exact distance travelled.

MATLAB **Note:**
Here is some MATLAB code to find a good approximation to the area under the curve $y = x^2$ over the domain $0 \le x \le 1$, using the left endpoints.

```
h=0.001; % the subinterval width
x=0:h:1-h; % the left subinterval endpoints (1000 values)
y=x.^2; % the height of each rectangle
area=sum(y*h) % the sum function just adds up each value of y*h
```

The answer to this is: area = 0.3328, which is already close to the exact answer 1/3. To get even closer, just take an even smaller value like $h =$ 0.0001. If you need to use right endpoints, just change the second code line to `x=h:h:1;` , ie x goes from $0 + h$ in steps of h to 1.

Activity

Stewart: sec 5.1 do exercises 11, 13, 15, 17.

Computer Activity

Stewart: sec 5.1 exercises 6, 7.

5.2 The Definite Integral

In this section the definite integral is introduced as the limit of so-called *Riemann sums*. This is crucial to your understanding of the integral and will help you greatly when you model real life applications which can be expressed as definite integrals.

Reading

Stewart: sec 5.2 study pp 354-364.
Take note of the definition of the definite integral $\int_a^b f(x)\,dx$ as the limit of a sum of function values multiplied by the step length Δx. **Note 1, 2,** and **3** are important. **Note 4** is interesting, but we will use equal length values of Δx. **Formulas 4,5,6)** on page 357 are interesting, but you will not be required to use them. Study **Examples 1,2,3,4** carefully. Note that the definite integral does not necessarily give the area under the curve. It is positive when the curve is above the x-axis and negative when the curve is below. The eight Properties of the Integral are important. By the way, Properties 2,3,4 indicate that the integral is a *linear operator*, ie

$$\int_a^b (C_1 f(x) + C_2 g(x))\,dx = C_1 \int_a^b f(x)\,dx + C_2 \int_a^b g(x)\,dx.$$

Remember that the derivative is also a linear operator.

MATLAB **Note:**
To change the MATLAB code for the *midpoint rule*, just use `x=h/2:h:1-h/2;` (see code in previous section), ie x goes from $h/2$ (the middle of the first subinterval) to $1-h/2$ (the middle of the last subinterval), in steps of h.

Activity

Stewart: sec 5.2 do exercises 1, 7, 11, 17, 27, 33(draw the graph), 39, 41, 43.

Computer Activity

Stewart: sec 5.2 do exercise 9 (use technology and increase n to something big to get more accuracy).

Note on the above Computer Activity: $\sqrt{x^3+1}$ in MATLAB is `sqrt(x.^3+1)`.

5.3 Evaluating Definite Integrals

Suppose you start your car at time $t = a$ at position $s(a)$, and drive in a straight line at velocity $v(t)$ until $t = b$ at position $s(b)$. Then the distance travelled is trivially $s(b) - s(a)$. But, as we have seen, the integral of the velocity is distance. Hence we have the very profound result

$$\int_a^b v(t)\,dt = s(b) - s(a).$$

But we know the $v(t) = \frac{ds}{dt}$, so we have

$$\int_a^b s'(t)\,dt = s(b) - s(a),$$

ie, the *total change in position* is the *integral of its derivative*. This illustrates the remarkable connection between integration and differentiation. In a sense, integration and differentiation are inverse operators, since one unravels the other.

Reading

Stewart: sec 5.3 study pp 366-373.
The Evaluation Theorem formalises the discussion above (where $s(t)$ is in fact the antiderivative of $s'(t)$). You can omit the proof of the Evaluation Theorem on page 367. Take particular note of Examples 1,2. A good notation for the antiderivative is $F(x) = \int f(x)\,dx$, and is called the *indefinite integral*. It is called indefinite, since it is a function, whereas the definite integral is a single number. The Table of Indefinite Integrals on page 369 should be familiar to you, since if you differentiate the right hand side, you get the function on the left. This Table becomes your working set of antiderivatives. Examples 3,4,5,6 are an excellent introduction to using the Table. The last subsection on Applications gives a very good indicator of various contexts for the integral.

Activity

Stewart: sec 5.3 do exercises 1, 3, 7, 9, 15, 23, 27, 37, 41, 47, 53, 57.

Computer Activity

Stewart: sec 5.3 do exercises 29, 31, 35.

5.4 The Fundamental Theorem of Calculus

In this section the two parts of the Fundamental Theorem of Calculus are formulated. A good way to remember each part is to *think*:

1. $\frac{d}{dx}\int f(x)\,dx = f(x)$;
2. $\int f'(x)\,dx = f(x)$.

Reading

Stewart: sec 5.4 study pp 377-382.
You are not required to learn the proof of the first part of the fundamental theorem, given in the last subsection.

We can use MATLAB to find a good approximation to the antiderivative of any reasonable, but complicated function. For example, if $f(x) = e^{x^2}$, we know from FTC1 that $F(x) = \int_0^x f(t)\,dt$ is an antiderivative of f. If we know that F starts at $F(0) = 3$ (say), we can then find the value of F at any other point x using

$$\int_0^x e^{t^2}\,dt = F(x) - F(0),$$

and so

$$F(x) = \int_0^x e^{t^2}\,dt + 3.$$

Then we use MATLAB to approximate the integral. Following is some code to find $F(2)$.

```
h=0.001; % chosen step size
t=0:h:2-h; % using left endpoints from t=0 to 2
f=exp(t.^2); % left values of f
F=sum(f*h)+3 % the integral + F(0)
```

This gives the result $F = F(2) = 19.4258$, a good approximation for the value of the antiderivative of e^{x^2} at $x = 2$, given that $F(0) = 3$.

Activity

Stewart: sec 5.4 do exercises 5, 7, 21.

Computer Activity

Do the following problem:
A car moves at velocity $v(t) = \sin(t^2) + te^t$. It starts at position $s(0) = 0$. Use technology to find a good approximation to its position $s(3)$ at $t = 3$. (Hint: you won't be able to find the antiderivative function for $\sin(t^2) + te^t$, so you will need to look at the MATLAB example above for guidance.)

5.5 The Substitution Rule

This section introduces a very handy technique for solving integrals of the form

$$\int_a^b f(g(x))g'(x)\,dx$$

by means of substitution. You will note that integrals of this form have a complicated function $f(g(x))$, but that they also have the luxury of the accompanying function $g'(x)$ as a partner. These integrals occur very frequently. You will remember from the chain rule that this type of combination of functions always result when you differentiate a composite function $f(g(x))$. So the method of substitution is really a process for unravelling the chain rule.

Reading

Stewart: sec 5.5 study pp 386-392.
Take note of the different types of substitution that can be used. It is not always obvious how to choose the substitution, so make sure that you work each of the Examples 1 to 8. The last subsection on Symmetry presents ways to sometimes speed up your calculations.

Activity

Stewart: sec 5.5 do exercises 1, 3, 5, 7, 13, 17, 27 (hint: write as $\sec^2 x \sec x \tan x$), 33, 41, 61.

5.6 Integration by Parts

Integration by parts is really a method for unravelling the derivative of a product. You should become very familiar with the equivalent formulas

$$\int uv'\,dx = uv - \int u'v\,dx$$

or,

$$\int u\,dv = uv - \int v\,du.$$

Reading

Stewart: sec 5.6 study pp 393-398.
Work Examples 1 and 2. Example 3 shows how you sometimes need to apply the integration by parts formula twice. Example 4 shows that you can sometimes use the cyclic behaviour for trigonometric functions to get there. Take note of Example 5 which finds the value of the integral and you can omit Example 6.

Activity

Stewart: sec 5.6 do exercises 1, 3, 7, 13, 15.

Computer Activity

Stewart: sec 5.6 do exercise 41 for the case $t = 5$ and check your answer by calculating a good approximation with MATLAB .

MATLAB **Note**: In the above Computer Activity use `v=t.^2.*exp(-t)` to get values of the function. See earlier MATLAB code for setting up the left endpoints and calculating the sum which gives the approximation to the definite integral.

5.7 Additional Techniques of Integration

There are many other special techniques for finding the integral. In this section we will only look at methods for trigonometric integrals.

Reading

Stewart: sec 5.7 study pp 400-402.
Omit the last subsection on Partial Fractions.
Example 1 is a good illustration of setting things up for a substitution. Example 2 is interesting if you know that $\sin^2 x = 0.5(1 - \cos 2x)$. In this course you don't need to learn that

$$\int \sin^2 x \, dx = 0.5(x - 0.5 \sin 2x) + C.$$

Example 3 is worth following since at last you are able to prove that the area of a circle is πr^2. Up until now you have taken this fact as an article of faith. By the way, Archimedes was able to find this over 2000 years ago. And he did not have the luxury of using the Fundamental Theorem. He might have known FTC1 or FTC2. We don't think he did, but he was one of the greatest thinkers of all time, so who knows!

Activity

Stewart: sec 5.7 do exercises 1 (hint: write as $\sin x \sin^2 x \cos^2 x$ and then use $\sin^2 x = 1 - \cos^2 x$), 7, 29 (hint: substitute $u = x^{0.5}$).

5.8 Integration Using Tables and Computer Algebra Systems

We will only have a quick look at this section. We do not use Computer Algebra Systems in this course, so you can omit that subsection.

Reading

Stewart: sec 5.8 study pp 405-407.
Omit the second half of page 407.
Work through Examples 2,3,4. Note that the table of integrals is on References pp 6-10 at the back of the book.
The rest of the section can be omitted.

Activity

Stewart: sec 5.8 do exercises 3, 5, 7.

5.9 Approximate Integration

In this section a number of methods are developed for approximating the value of definite integrals. We will see them in a later mathematics course, so we will only do the first page of this section on some, by now, familiar methods for approximating the integral. We have already seen how to do the approximation using left and right endpoints and also for the midpoint rule.

Reading

Stewart: sec 5.9 study page 412 only.

Here is a reminder of the MATLAB code for approximating $\int_0^3 e^{x^2}\,dx$:

Left Endpoint Approximation:

```
h=0.001;  % your choice of step length
x=0:h:3-h; % the left hand endpoints
y=exp(x.^2); % the values at the left endpoints
left=sum(y*h) % the answer
```

Right Endpoint Approximation:

```
h=0.001;  % your choice of step length
x=h:h:3; % the right hand endpoints
y=exp(x.^2); % the values at the right endpoints
right=sum(y*h) % the answer
```

Midpoint Approximation:

```
h=0.001;  % your choice of step length
x=h/2:h:3-h/2; % the midpoints
y=exp(x.^2); % the values at the midpoints
mid=sum(y*h) % the answer
```

> **Activity**
>
> Stewart: sec 5.9 do exercise 5(a). Note: if you want 8 midpoints, you will need to firstly calculate the value of h. Here it will be $h = \frac{\pi - 0}{8}$. In MATLAB that is h=(pi-0)/8.

Chapter Summary

The last section on Improper Integrals may be omitted, but should be read by mathematics majors.

Standard methods of integration have been introduced in this chapter. There is much useful material here. You are now equipped with the foundation tools of both differential and integral calculus. This will form the basis of your abilities as a technician of change. If it moves, you can now hit it with calculus. The next, and final chapter in this course is the chance to apply some of these skills to more practical situations.

Before moving on, try some of the Review questions at the end of this chapter.

Chapter 6

Applications of Integration

INTRODUCTION

In this, the final chapter, we study applications of the definite integral including finding areas, volumes, averages, as well as some applications to physics and engineering. We omit some sections and subsections as described following.

6.1 More about Areas

So far we have studied areas under curves. Using similar methods, it is easy now to extend the technique to finding areas between curves.

Reading

Stewart: sec 6.1 study pp 441-445.
Omit the last subsection on Areas Enclosed by Parametric Curves.

Activity

Stewart: sec 6.1 do exercises 1, 7, 9, 25.

Computer Activity

Stewart: sec 6.1 do exercise 17.

6.2 Volumes

In this section we use Riemann sums to find volumes of revolution as well as volumes whose cross-sections are known.

Reading

Stewart: sec 6.1 study pp 447-456, but omit the last subsection on Cylindrical Shells.
For volumes of revolution, we will only use the disc method. The trick here is to find the volume of a disc section of thickness Δx and radius given by $y = f(x)$. The volume of the disc is then $\pi y^2 \Delta x = \pi f(x)^2 \Delta x$. Then sum all the discs to get $\sum \pi f(x_i)^2 \Delta x$. This is just a Riemann sum and in the limit as $\Delta x \to 0$ we get the definite integral $\int_a^b \pi f(x)^2 \, dx$ which can be evaluated (or approximated) by the methods of the last chapter.

Activity

Stewart: sec 6.1 do exercises 1, 3, 5, 11*.

6.3 Arc Length

Omit this section which will be studied in the following course MAT2100.

6.4 Average Value of a Function

The average value of a function $f(x)$, over some domain $a \leq x \leq b$, is quite easy to evaluate. It is just

$$\text{average } = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

Reading

Stewart: sec 6.4 study pp 467-469. You may omit the Mean Value Theorem for Integrals and omit Example 2.

Activity

Stewart: sec 6.4 do exercises 1, 5, 15.

Computer Activity

Stewart: sec 6.4 do exercise 7.

6.5 Applications to Physics and Engineering

Three applications are included in this chapter, Work, Hydrostatic Pressure & Force, and Centres of Mass.

Reading

Stewart: sec 6.5 study pp 471-479.
Study the three application subsections and work through Examples 1-7. Further notes follow.

Non-uniform density
In the analysis of this section it has been assumed that the density of the material or the medium is constant. How do we handle things if the density is not constant? Here is an example which follows the modelling procedure, develops the integral, and then we solve it.

Example
Consider a rod from $x = 0$ to $x = 4$ whose density at a point x is given by $\rho(x) = 2 + \frac{1}{x+1}$ kilograms per metre. Find the mass of the rod.

Solution:
Split the rod up into n subintervals of length Δx. On a typical subinterval (ie an element of length Δx) we then assume the density to have its value at a point x_i on the element (eg the left endpoint) and thus be constant across the element with value $\rho(x_i)$. The mass of the element is then just (density)×(length), or $\rho(x_i)\Delta x$. Summing all of the elements gives $\sum \rho(x_i)\Delta x$. This is a left endpoint Riemann sum, and in the limit as $\Delta x \to 0$, we get $\int_0^4 \rho(x)\,dx$.

Thus the mass of the rod is $\int_0^4 \left(2 + \frac{1}{x+1}\right) dx$. Evaluating

$$\begin{aligned}
\int_0^4 \left(2 + \frac{1}{x+1}\right) dx &= [2x + \ln(x+1)]_0^4 \\
&= (8 + \ln 5) - (0 + \ln 1) \\
&\approx (8 + 1.6094) - (0 + 0) \\
&= 9.6094\text{kg}.
\end{aligned}$$

Activity

Stewart: sec 6.5 do exercises 1, 15, 23, 35.

Chapter Summary The remainder of this chapter is to be omitted, but if you are majoring in business, biology, or mathematics, you should read selections of Sections 6.6 or 6.7.

We have covered a lot of ground in this course and now it is time to study for the exam. Good luck and hope to see you in the following course MAT2100.

III

Appendix A and B: Readings for Module I

Appendix A: Vectors

Appendix A contains readings on Vectors for Chapter 0 of Module 1 in this Study Guide.

The readings are an excerpt from *Elementary Linear Algebra: International Edition*, by Grossman, published by Saunders College Publishing, pp 154–189, 192–207. Solutions to odd-numbered exercises are offered in A33–37 and A39.

Additional Summary Sheets

Separating Appendix A and B are additional summaries intended to serve as further review of Chapters 1, 2 and 3 of Module I, Systems of Linear Equations, Matrices and Determinants. Some are excerpts from Grossman's *Elementary Linear Algebra*, pp 151–152.

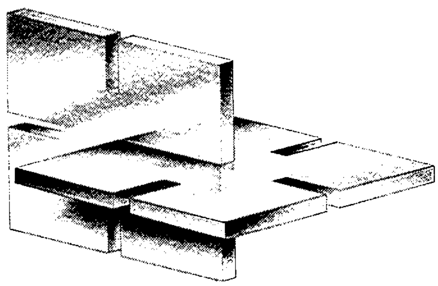
Appendix B: Complex Numbers

Appendix B contains the readings for Complex Numbers, Chapter 4 of Module 1 in this Study Guide, and solutions to the odd-numbered exercises.

They are excerpts from *Elementary Linear Algebra: Third Edition*, by Larson & Edwards, published by Heath, pp 431–473 and A69–76.

Appendix A

Grossman, SI 1991, *Elementary linear algebra*, International edn, Saunders College Publishing, New York, pp. 154–189, 192–207, A33–7 & A39.



3

Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$

3.1 VECTORS IN THE PLANE

As we defined it in Section 1.5, $\mathbb{R}^2$ is the set of vectors (x_1, x_2) with x_1 and x_2 real numbers. Since any point in the plane can be written in the form (x, y), it is apparent that any point in the plane can be thought of as a vector in $\mathbb{R}^2$, and vice versa. Thus the terms "the plane" and "$\mathbb{R}^2$" are often used interchangeably. However, for a variety of physical applications (including the notions of force, velocity, acceleration, and momentum), it is important to think of a vector not as a point but as an entity having "length" and "direction." Now we shall see how this is done.

Directed line segment

Let P and Q be two points in the plane. Then the **directed line segment** from P to Q, denoted by $\overrightarrow{PQ}$, is the straight-line segment that extends from P to Q (see

Figure 3.1*a*). Note that the directed line segments $\overrightarrow{PQ}$ and $\overrightarrow{QP}$ are different since they point in opposite directions (Figure 3.1*b*).

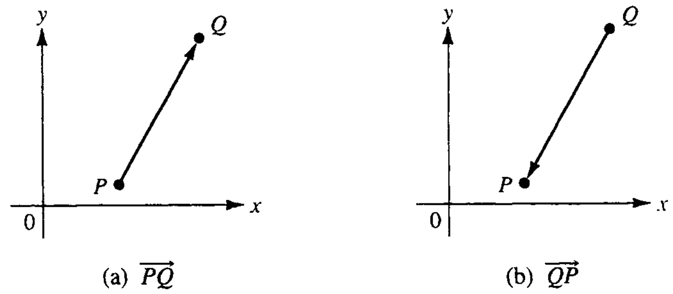


Figure 3.1 The directed line segments $\overrightarrow{PQ}$ and $\overrightarrow{QP}$ point in opposite directions

Initial point

Terminal point

Equivalent directed line segment

The point P in the directed line segment $\overrightarrow{PQ}$ is called the **initial point** of the segment and the point Q is called the **terminal point.** The two major properties of a directed line segment are its magnitude (length) and its direction. If two directed line segments $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ have the same magnitude and direction, we say that they are **equivalent** no matter where they are located with respect to the origin. The directed line segments in Figure 3.2 are all equivalent.

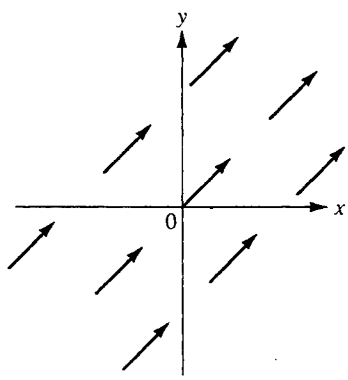


Figure 3.2 A set of equivalent directed line segments

DEFINITION 1 **Geometric Definition of a Vector** The set of all directed line segments equivalent to a given directed line segment is called a **vector.** Any directed line segment in that set is called a **representation** of the vector.

Remark. The directed line segments in Figure 3.2 are all representations of the same vector.

From Definition 1 we see that a given vector $\mathbf{v}$ can be represented in many different ways. Let $\overrightarrow{PQ}$ be a representation of $\mathbf{v}$. Then, without changing magnitude or direction, we can move $\overrightarrow{PQ}$ in a parallel way so that its initial point is shifted to the origin. We then obtain the directed line segment $\overrightarrow{0R}$, which is another representation of the vector $\mathbf{v}$ (see Figure 3.3). Now suppose that R has the Cartesian

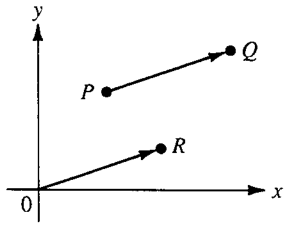


Figure 3.3 We can move $\overrightarrow{PQ}$ to obtain an equivalent directed line segment with its initial point at the origin. Note that $\overrightarrow{0R}$ and $\overrightarrow{PQ}$ are parallel and have the same length

coordinates (a, b). Then we can describe the directed line segment $\overrightarrow{0R}$ by the coordinates (a, b). That is, $\overrightarrow{0R}$ is the directed line segment with initial point $(0, 0)$ and terminal point (a, b). Since one representation of a vector is as good as another, we can write the vector $\mathbf{v}$ as (a, b).

DEFINITION 2 **Algebraic Definition of a Vector** A **vector** $\mathbf{v}$ in the xy-plane is an ordered pair of real numbers (a, b). The numbers a and b are called the **components** of the vector $\mathbf{v}$. The **zero vector** is the vector $(0, 0)$.

Remark 1. With this definition, a point in the xy-plane can be thought of as a vector originating at the origin and terminating at that point.

Remark 2. The zero vector has a magnitude of zero. Therefore, since the initial and terminal points coincide, we say that the zero vector has *no direction*.

Remark 3. We emphasize that Definitions 1 and 2 describe precisely the same objects. Each point of view (geometric and algebraic) has its advantages. Definition 2 is the definition of a 2-vector that we have been using all along.

Magnitude or length of a vector

Since a vector is really a set of equivalent line segments, we define the **magnitude** or **length** of a vector as the length of any one of its representations and its **direction** as the direction of any one of its representations. Using the representation $\overrightarrow{0R}$ and writing the vector $\mathbf{v} = (a, b)$, we find that

$$\boxed{|\mathbf{v}| = \text{magnitude of } \mathbf{v} = \sqrt{a^2 + b^2}} \tag{1}$$

This follows from the Pythagorean theorem (see Figure 3.4). We have used the notation $|\mathbf{v}|$ to denote the magnitude of $\mathbf{v}$. Note that $|\mathbf{v}|$ is a *scalar*.

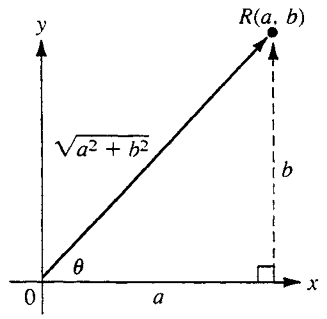


Figure 3.4 The magnitude of a vector with x-coordinate a and y-coordinate b is $\sqrt{a^2 + b^2}$

EXAMPLE 1 **Calculating the Magnitudes of Six Vectors** Calculate the magnitudes of the vectors **(i)** $(2, 2)$; **(ii)** $(2, 2\sqrt{3})$; **(iii)** $(-2\sqrt{3}, 2)$; **(iv)** $(-3, -3)$; **(v)** $(6, -6)$; **(vi)** $(0, 3)$.

Solution

i. $|\mathbf{v}| = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2}$

ii. $|\mathbf{v}| = \sqrt{2^2 + (2\sqrt{3})^2} = 4$

iii. $|\mathbf{v}| = \sqrt{(-2\sqrt{3})^2 + 2^2} = 4$

iv. $|\mathbf{v}| = \sqrt{(-3)^2 + (-3)^2} = \sqrt{18} = 3\sqrt{2}$

v. $|\mathbf{v}| = \sqrt{6^2 + (-6)^2} = \sqrt{72} = 6\sqrt{2}$

vi. $|\mathbf{v}| = \sqrt{0^2 + 3^2} = \sqrt{9} = 3$ ■

Direction of a vector

We now define the **direction** of the vector $\mathbf{v} = (a, b)$ to be the angle θ, measured in radians, that the vector makes with the positive x-axis. By convention, we choose θ such that $0 \leq \theta < 2\pi$. It follows from Figure 3.4 that if $a \neq 0$, then

$$\boxed{\tan \theta = \frac{b}{a}} \qquad (2)$$

Note: Tan θ is periodic of period π, so if $a \neq 0$ there are always *two* numbers in $[0, 2\pi)$ such that $\tan \theta = b/a$. For example, $\tan \pi/4 = \tan 5\pi/4 = 1$. In order to determine θ uniquely, we need to determine the quadrant of $\mathbf{v}$ as we shall see in the next example.

EXAMPLE 2 **Calculating the Directions of Six Vectors** Calculate the directions of the vectors in Example 1.

Solution These six vectors are depicted in Figure 3.5.

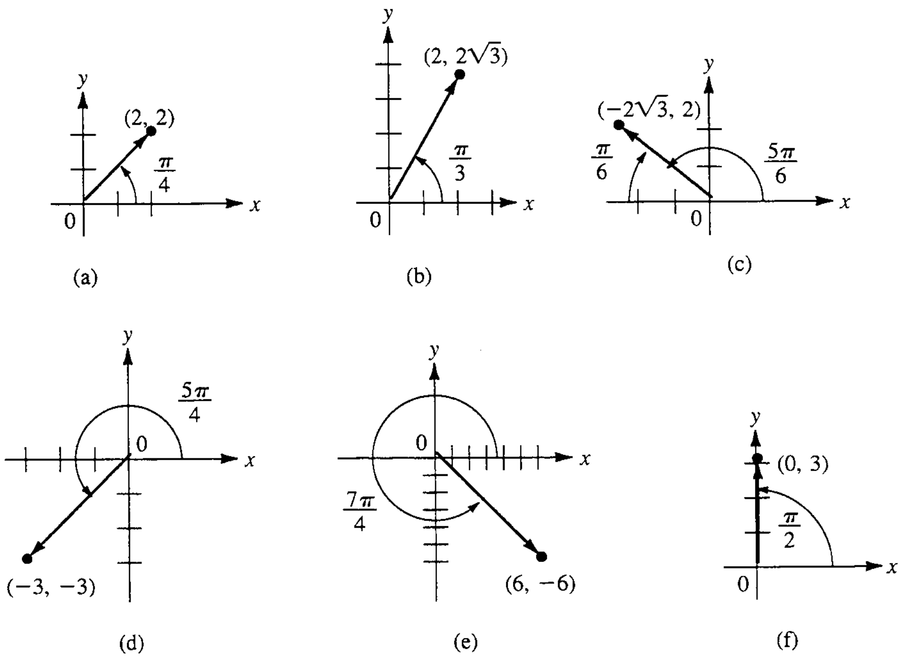


Figure 3.5 The directions of six vectors

i. Here $\mathbf{v}$ is in the first quadrant and since $\tan\theta = 2/2 = 1$, $\theta = \pi/4$.

ii. Here $\theta = \tan^{-1} 2\sqrt{3}/2 = \tan^{-1}\sqrt{3} = \pi/3$ (since $\mathbf{v}$ is in the first quadrant).

iii. We see that $\mathbf{v}$ is in the second quadrant and, since $\tan^{-1} 2/2\sqrt{3} = \tan^{-1} 1/\sqrt{3} = \pi/6$, we see from Figure 3.5$c$ that $\theta = \pi - (\pi/6) = 5\pi/6$.

iv. Here $\mathbf{v}$ is in the third quadrant and, since $\tan^{-1} 1 = \pi/4$, we find that $\theta = \pi + (\pi/4) = 5\pi/4$.

v. Since $\mathbf{v}$ is in the fourth quadrant and $\tan^{-1}(-1) = -\pi/4$, we get $\theta = 2\pi - (\pi/4) = 7\pi/4$.

vi. We cannot use equation (2) because b/a is undefined. However, we see in Figure 5f that $\theta = \pi/2$. ■

In general, if $b > 0$

$$\boxed{\text{Direction of } (0, b) = \frac{\pi}{2} \quad \text{and} \quad \text{direction of } (0, -b) = \frac{3\pi}{2} \qquad b > 0}$$

In Section 1.5 we defined vector addition and scalar multiplication. What do these concepts mean geometrically? We start with scalar multiplication. If $\mathbf{v} = (a, b)$, then $\alpha\mathbf{v} = (\alpha a, \alpha b)$. We find that

$$|\alpha\mathbf{v}| = \sqrt{\alpha^2 a^2 + \alpha^2 b^2} = |\alpha|\sqrt{a^2 + b^2} = |\alpha|\,|\mathbf{v}| \tag{3}$$

That is,

> Multiplying a vector by a nonzero scalar has the effect of multiplying the length of the vector by the absolute value of that scalar.

Moreover, if $\alpha > 0$, then $\alpha\mathbf{v}$ is in the same quadrant as $\mathbf{v}$, and therefore the direction of $\alpha\mathbf{v}$ is the *same* as the direction of $\mathbf{v}$ since $\tan^{-1}(\alpha b/\alpha a) = \tan^{-1}(b/a)$. If $\alpha < 0$, then $\alpha\mathbf{v}$ points in the direction opposite to that of $\mathbf{v}$. In other words,

$$\boxed{\begin{array}{l}\text{Direction of } \alpha\mathbf{v} = \text{direction of } \mathbf{v}, \text{ if } \alpha > 0\\ \text{Direction of } \alpha\mathbf{v} = \text{direction of } \mathbf{v} + \pi, \text{ if } \alpha < 0\end{array}} \tag{4}$$

EXAMPLE 3 **Multiplying a Vector by a Scalar** Let $\mathbf{v} = (1, 1)$. Then $|\mathbf{v}| = \sqrt{1 + 1} = \sqrt{2}$ and $|2\mathbf{v}| = |(2, 2)| = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2} = 2|\mathbf{v}|$. Further, $|-2\mathbf{v}| = \sqrt{(-2)^2 + (-2)^2} = 2\sqrt{2} = 2|\mathbf{v}|$. Moreover, the direction of $2\mathbf{v}$ is $\pi/4$, whereas

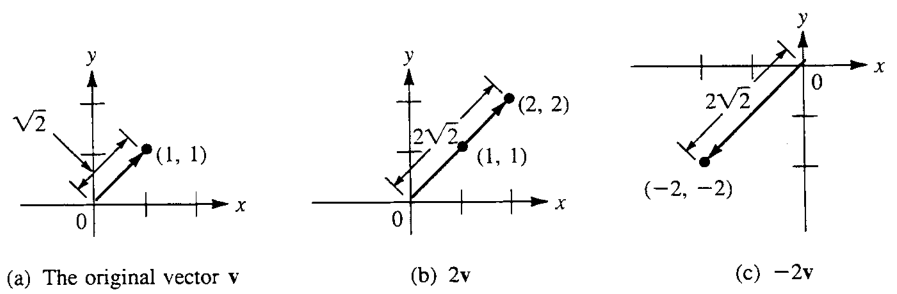


Figure 3.6 The vector $2\mathbf{v}$ has the same direction as $\mathbf{v}$ and twice its magnitude. The vector $-2\mathbf{v}$ has the opposite direction of $\mathbf{v}$ and twice its magnitude

the direction of $-2\mathbf{v}$ is $5\pi/4$ (see Figure 3.6). ■

Now suppose we add the vectors $\mathbf{u} = (a_1, b_1)$ and $\mathbf{v} = (a_2, b_2)$ as in Figure 3.7. From the figure we see that the vector $\mathbf{u} + \mathbf{v} = (a_1 + a_2, b_1 + b_2)$ can be

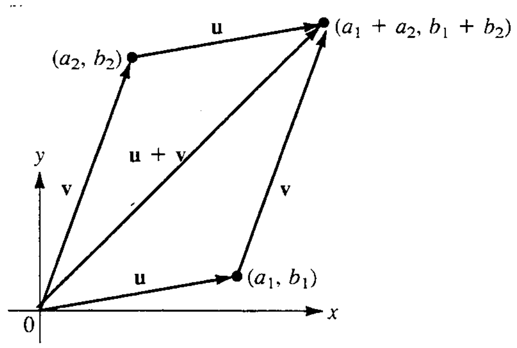


Figure 3.7 The parallelogram rule for adding vectors

obtained by shifting the representation of the vector **v** so that its initial point coincides with the terminal point (a_1, b_1) of the vector **u**. We can therefore obtain the vector **u** + **v** by drawing a parallelogram with one vertex at the origin and sides **u** and **v**. Then **u** + **v** is the vector that points from the origin along the diagonal of the parallelogram.

Note. Since a straight line is the shortest distance between two points, it immediately follows from Figure 3.7 that

Triangle Inequality

$$|\mathbf{u} + \mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}| \tag{5}$$

For reasons obvious from Figure 3.7, inequality (5) is called the **triangle inequality.**

We can also use Figure 3.7 to obtain a geometric representation of the vector **u** − **v**. Since **u** = **u** − **v** + **v**, the vector **u** − **v** is the vector that must be added to **v** to obtain **u**. This fact is illustrated in Figure 3.8*a*. A similar fact is illustrated in Figure 3.8*b*.

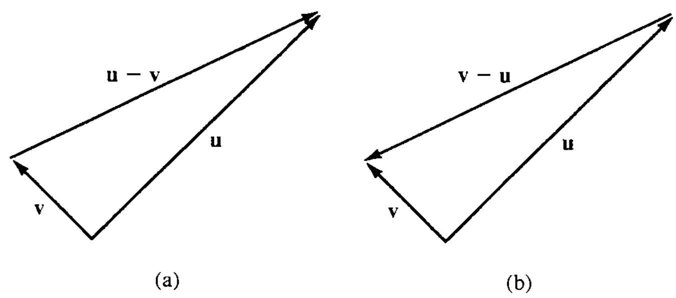


Figure 3.8 The vectors **u** − **v** and **v** − **u** have the same magnitude but point in opposite directions

There are two special vectors in $\mathbb{R}^2$ that allow us to represent other vectors in $\mathbb{R}^2$ in a convenient way. We denote the vector $(1, 0)$ by the symbol $\mathbf{i}$ and the vector $(0, 1)$ by the vector $\mathbf{j}$. (See Figure 3.9.) If $\mathbf{v} = (a, b)$ is any vector in the plane, then

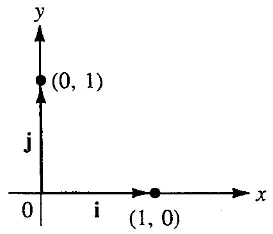


Figure 3.9 The vectors **i** and **j**

since $(a, b) = a(1, 0) + b(0, 1)$, we may write

$$\boxed{\mathbf{v} = (a, b) = a\mathbf{i} + b\mathbf{j}} \tag{6}$$

With this representation we say that $\mathbf{v}$ is *resolved into its horizontal and vertical components*. The vectors $\mathbf{i}$ and $\mathbf{j}$ have two properties:

i. Neither one is a multiple of the other. (In the terminology of Chapter 4, they are *linearly independent*.)

ii. Any vector $\mathbf{v}$ can be written in terms of $\mathbf{i}$ and $\mathbf{j}$ as in equation (6).†

Historical Note. The symbols $\mathbf{i}$ and $\mathbf{j}$ were first used by Hamilton. He defined his quaternion as a quantity of the form $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, where a is the "scalar part" and $b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ the "vector part." In Section 3.3 we shall write vectors in space in the form $b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$.

Under these two conditions $\mathbf{i}$ and $\mathbf{j}$ are said to form a **basis** in $\mathbb{R}^2$. We shall discuss bases in arbitrary vector spaces in Chapter 4.

We now define a kind of a vector that is very useful in certain applications.

DEFINITION 3 **Unit Vector** A **unit vector u** is a vector that has length 1.

EXAMPLE 4 **A Unit Vector** The vector $\mathbf{u} = (1/2)\mathbf{i} + (\sqrt{3}/2)\mathbf{j}$ is a unit vector since

$$|\mathbf{u}| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{3}{4}} = 1$$

■

† In equation (6) we say that $\mathbf{v}$ can be written as a *linear combination* of $\mathbf{i}$ and $\mathbf{j}$. We shall discuss the notion of linear combination in Section 4.5.

Let $\mathbf{u} = a\mathbf{i} + b\mathbf{j}$ be a unit vector. Then $|\mathbf{u}| = \sqrt{a^2 + b^2} = 1$, so that $a^2 + b^2 = 1$ and $\mathbf{u}$ can be represented by a point on the unit circle (see Figure 3.10).

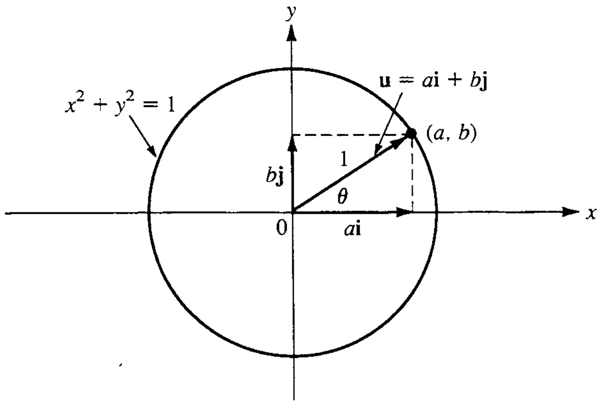


Figure 3.10 The terminal point of a unit vector with initial point at the origin lies on the unit circle

If θ is the direction of $\mathbf{u}$, then we immediately see that $a = \cos\theta$ and $b = \sin\theta$. Thus any unit vector $\mathbf{u}$ can be written in the form

$$\boxed{\mathbf{u} = (\cos\theta)\mathbf{i} + (\sin\theta)\mathbf{j}} \tag{7}$$

where θ is the direction of $\mathbf{u}$.

EXAMPLE 5 **Writing a Unit Vector as $(\cos\theta)\mathbf{i} + (\sin\theta)\mathbf{j}$** The unit vector $\mathbf{u} = (1/2)\mathbf{i} + (\sqrt{3}/2)\mathbf{j}$ of Example 4 can be written in the form of (7) with $\theta = \cos^{-1}(1/2) = \pi/3$. ■

We also have (see Problem 17).

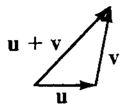
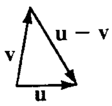
> Let $\mathbf{v}$ be any nonzero vector. Then $\mathbf{u} = \mathbf{v}/|\mathbf{v}|$ is a unit vector having the same direction as $\mathbf{v}$.

EXAMPLE 6 **Finding a Unit Vector Having the Same Direction as a Given Vector** Find a unit vector having the same direction as $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j}$.

Solution Here $|\mathbf{v}| = \sqrt{4 + 9} = \sqrt{13}$, so $\mathbf{u} = \mathbf{v}/|\mathbf{v}| = (2/\sqrt{13})\mathbf{i} - (3/\sqrt{13})\mathbf{j}$ is the required unit vector. ■

We conclude this section with a summary of the properties of vectors (Table 3.1).

Table 3.1

Object	Intuitive Definition	Expression in terms of components if $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$, and $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2)$
Vector $\mathbf{v}$	An object having magnitude and direction	$v_1\mathbf{i} + v_2\mathbf{j}$ or (v_1, v_2)
$\lvert\mathbf{v}\rvert$	Magnitude (or length) of $\mathbf{v}$	$\sqrt{v_1^2 + v_2^2}$
$\alpha\mathbf{v}$	$\nearrow\mathbf{v}$ $\nearrow\alpha\mathbf{v}$ (In this sketch $\alpha = 2$)	$\alpha v_1\mathbf{i} + \alpha v_2\mathbf{j}$ or $(\alpha v_1, \alpha v_2)$
$-\mathbf{v}$	$\nearrow\mathbf{v}$ $\swarrow -\mathbf{v}$	$-v_1\mathbf{i} - v_2\mathbf{j}$ or $(-v_1 - v_2)$ or $-(v_1, v_2)$
$\mathbf{u} + \mathbf{v}$		$(u_1 + v_1)\mathbf{i} + (u_2 + v_2)\mathbf{j}$ or $(u_1 + v_1, u_2 + v_2)$
$\mathbf{u} - \mathbf{v}$		$(u_1 - v_1)\mathbf{i} + (u_2 - v_2)\mathbf{j}$ or $(u_1 - v_1, u_2 - v_2)$

PROBLEMS 3.1

In Problems 1–12 find the magnitude and direction of the given vector.

1. $\mathbf{v} = (4, 4)$
2. $\mathbf{v} = (-4, 4)$
3. $\mathbf{v} = (4, -4)$
4. $\mathbf{v} = (-4, -4)$
5. $\mathbf{v} = (\sqrt{3}, 1)$
6. $\mathbf{v} = (1, \sqrt{3})$
7. $\mathbf{v} = (-1, \sqrt{3})$
8. $\mathbf{v} = (1, -\sqrt{3})$
9. $\mathbf{v} = (-1, -\sqrt{3})$
10. $\mathbf{v} = (1, 2)$
11. $\mathbf{v} = (-5, 8)$
12. $\mathbf{v} = (11, -14)$
13. Let $\mathbf{u} = (2, 3)$ and $\mathbf{v} = (-5, 4)$. Find: **(a)** $3\mathbf{u}$; **(b)** $\mathbf{u} + \mathbf{v}$; **(c)** $\mathbf{v} - \mathbf{u}$; **(d)** $2\mathbf{u} - 7\mathbf{v}$. Sketch these vectors.
14. Let $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j}$ and $\mathbf{v} = -4\mathbf{i} + 6\mathbf{j}$. Find: (a) $\mathbf{u} + \mathbf{v}$; (b) $\mathbf{u} - \mathbf{v}$; (c) $3\mathbf{u}$; (d) $-7\mathbf{v}$; (e) $8\mathbf{u} - 3\mathbf{v}$; (f) $4\mathbf{v} - 6\mathbf{u}$. Sketch these vectors.
15. Show that the vectors $\mathbf{i}$ and $\mathbf{j}$ are unit vectors.
16. Show that the vector $(1/\sqrt{2})\mathbf{i} + (1/\sqrt{2})\mathbf{j}$ is a unit vector.
17. Show that if $\mathbf{v} = a\mathbf{i} + b\mathbf{j} \neq \mathbf{0}$, then $\mathbf{u} = (a/\sqrt{a^2 + b^2})\mathbf{i} + (b/\sqrt{a^2 + b^2})\mathbf{j}$ is a unit vector having the same direction as $\mathbf{v}$.

In Problems 18–21 find a unit vector having the same direction as the given vector.

18. $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$

19. $\mathbf{v} = \mathbf{i} - \mathbf{j}$

20. $\mathbf{v} = -3\mathbf{i} + 4\mathbf{j}$

21. $\mathbf{v} = a\mathbf{i} + a\mathbf{j}$; $a \neq 0$.

22. If $\mathbf{v} = a\mathbf{i} + b\mathbf{j}$, show that $a/\sqrt{a^2+b^2} = \cos\theta$ and $b/\sqrt{a^2+b^2} = \sin\theta$, where θ is the direction of $\mathbf{v}$.

23. If $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j}$, find $\sin\theta$ and $\cos\theta$.

24. If $\mathbf{v} = -3\mathbf{i} + 8\mathbf{j}$, find $\sin\theta$ and $\cos\theta$.

A vector $\mathbf{v}$ has a direction opposite to that of a vector $\mathbf{u}$ if direction $\mathbf{v}$ = direction $\mathbf{u} + \pi$. In Problems 25–28 find a unit vector $\mathbf{v}$ that has a direction opposite the direction of the given vector $\mathbf{u}$.

25. $\mathbf{u} = \mathbf{i} + \mathbf{j}$

26. $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j}$

27. $\mathbf{u} = -3\mathbf{i} + 4\mathbf{j}$

28. $\mathbf{u} = -2\mathbf{i} + 3\mathbf{j}$

29. Let $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j}$ and $\mathbf{v} = -\mathbf{i} + 2\mathbf{j}$. Find a unit vector having the same direction as: **(a)** $\mathbf{u} + \mathbf{v}$; **(b)** $2\mathbf{u} - 3\mathbf{v}$; **(c)** $3\mathbf{u} + 8\mathbf{v}$.

30. Let $P = (c, d)$ and $Q = (c + a, d + b)$. Show that the magnitude of $\overrightarrow{PQ}$ is $\sqrt{a^2+b^2}$.

31. Show that the direction of $\overrightarrow{PQ}$ in Problem 30 is the same as the direction of the vector (a, b). [*Hint:* If $R = (a, b)$, show that the line passing through the points P and Q is parallel to the line passing through the points 0 and R.]

In Problems 32–35 find a vector $\mathbf{v}$ having the given magnitude and direction.

32. $|\mathbf{v}| = 3$; $\theta = \pi/6$

33. $|\mathbf{v}| = 8$; $\theta = \pi/3$

34. $|\mathbf{v}| = 1$; $\theta = \pi/4$

35. $|\mathbf{v}| = 6$; $\theta = 2\pi/3$.

***36.** Show algebraically (that is, strictly from the definitions of vector addition and magnitude) that for any two vectors $\mathbf{u}$ and $\mathbf{v}$, $|\mathbf{u} + \mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}|$.

37. Show that if neither $\mathbf{u}$ nor $\mathbf{v}$ is the zero vector, then $|\mathbf{u} + \mathbf{v}| = |\mathbf{u}| + |\mathbf{v}|$ if and only if $\mathbf{u}$ is a positive scalar multiple of $\mathbf{v}$.

3.2 THE SCALAR PRODUCT AND PROJECTIONS IN $\mathbb{R}^2$

In Section 1.7 we defined the scalar product of two vectors. If $\mathbf{u} = (a_1, b_1)$ and $\mathbf{v} = (a_2, b_2)$, then

$$\boxed{\mathbf{u}\cdot\mathbf{v} = a_1a_2 + b_1b_2} \tag{1}$$

We now see how the scalar product can be interpreted geometrically.

DEFINITION 1 **Angle Between Vectors** Let $\mathbf{u}$ and $\mathbf{v}$ be two nonzero vectors. Then the **angle** φ **between u and v** is defined to be the smallest nonnegative angle† between the

†This angle will be in the interval $[0, \pi]$.

representations of **u** and **v** that have the origin as their initial points. If $\mathbf{u} = \alpha\mathbf{v}$ for some scalar α, then we define $\varphi = 0$ if $\alpha > 0$ and $\varphi = \pi$ if $\alpha < 0$.

This definition is illustrated in Figure 3.11. Note that φ can always be chosen to be a nonnegative angle in the interval $[0, \pi]$.

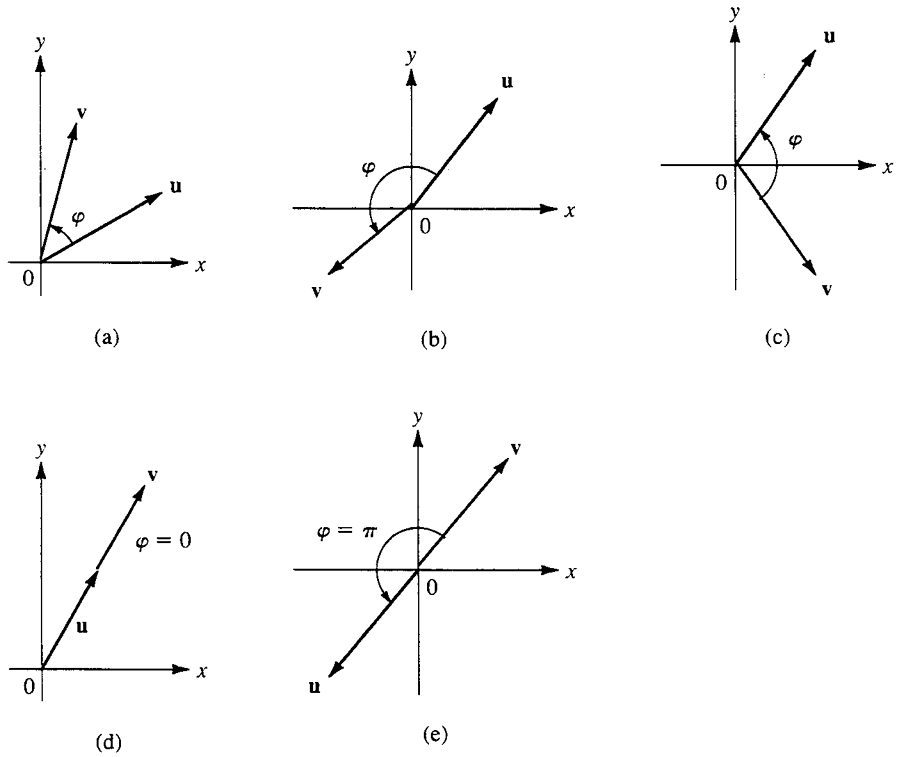


Figure 3.11 The angle φ between two vectors

THEOREM 1 Let **v** be a vector. Then

$$\boxed{|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v}} \tag{2}$$

Proof Let $\mathbf{v} = (a, b)$. Then

$$|\mathbf{v}|^2 = a^2 + b^2$$

and

$$\mathbf{v} \cdot \mathbf{v} = (a, b) \cdot (a, b) = a \cdot a + b \cdot b = a^2 + b^2 = |\mathbf{v}|^2 \qquad \blacksquare$$

THEOREM 2 Let $\mathbf{u}$ and $\mathbf{v}$ be two nonzero vectors. If φ is the angle between them, then

$$\boxed{\cos\varphi = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|}} \tag{3}$$

Proof The law of cosines (see Problem 2.5.10, page 150) states that in the triangle of Figure 3.12

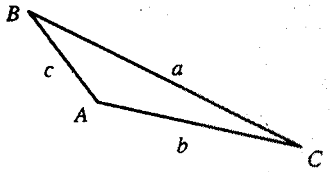


Figure 3.12 A triangle with sides a, b, and c

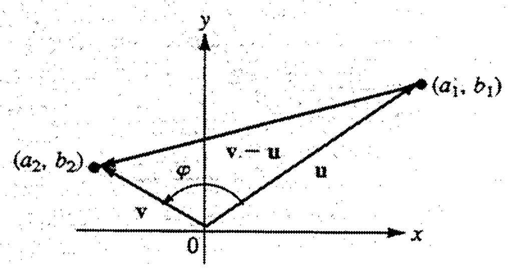


Figure 3.13 A triangle with sides $|\mathbf{u}|$, $|\mathbf{v}|$, and $|\mathbf{v}-\mathbf{u}|$

$$c^2 = a^2 + b^2 - 2ab\cos C$$

We now place the representations of $\mathbf{u}$ and $\mathbf{v}$ with initial points at the origin so that $\mathbf{u} = (a_1, b_1)$ and $\mathbf{v} = (a_2, b_2)$ (see Figure 3.13). Then, from the law of cosines, $|\mathbf{v}-\mathbf{u}|^2 = |\mathbf{v}|^2 + |\mathbf{u}|^2 - 2|\mathbf{u}|\,|\mathbf{v}|\cos\varphi$. But

$$|\mathbf{v}-\mathbf{u}|^2 \overset{\text{from (2)}}{=} (\mathbf{v}-\mathbf{u})\cdot(\mathbf{v}-\mathbf{u}) \overset{\text{Theorem 1 } (iii) \text{ on page 44}}{=} \mathbf{v}\cdot\mathbf{v} - 2\mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{u}$$
$$= |\mathbf{v}|^2 - 2\mathbf{u}\cdot\mathbf{v} + |\mathbf{u}|^2$$

Thus, after simplification, we obtain $-2\mathbf{u}\cdot\mathbf{v} = -2|\mathbf{u}|\,|\mathbf{v}|\cos\varphi$, from which the theorem follows. ■

Remark. Using Theorem 1 we could define the scalar product $\mathbf{u}\cdot\mathbf{v}$ by

$$\boxed{\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}|\,|\mathbf{v}|\cos\varphi}$$

EXAMPLE 1 **Computing the Angle Between Two Vectors** Find the angle between the vectors $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j}$ and $\mathbf{v} = -7\mathbf{i} + \mathbf{j}$.

Solution $\mathbf{u} \cdot \mathbf{v} = -14 + 3 = -11$, $|\mathbf{u}| = \sqrt{2^2 + 3^2} = \sqrt{13}$, and $|\mathbf{v}| = \sqrt{(-7)^2 + 1^2} = \sqrt{50}$. Hence

$$\cos \varphi = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{-11}{\sqrt{13}\,\sqrt{50}} = \frac{-11}{\sqrt{650}} \approx -0.431455497\dagger$$

so

$$\varphi = \cos^{-1}(-0.431455497) \approx 2.0169\ddagger \;(\approx 115.6°) \qquad \blacksquare$$

DEFINITION 2 **Parallel Vectors** Two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ are **parallel** if the angle between them is zero or π. Note that parallel vectors have the same or opposite directions.

EXAMPLE 2 **Two Parallel Vectors** Show that the vectors $\mathbf{u} = (2, -3)$ and $\mathbf{v} = (-4, 6)$ are parallel.

Solution

$$\cos \varphi = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{-8 - 18}{\sqrt{13}\,\sqrt{52}} = \frac{-26}{\sqrt{13}(2\sqrt{13})} = \frac{-26}{2(13)} = -1$$

Hence $\varphi = \pi$ (so that $\mathbf{u}$ and $\mathbf{v}$ have opposite directions). $\blacksquare$

THEOREM 3 If $\mathbf{u} \neq \mathbf{0}$, then $\mathbf{v} = \alpha\mathbf{u}$ for some nonzero constant α if and only if $\mathbf{u}$ and $\mathbf{v}$ are parallel.

Proof The proof is left as an exercise (see Problem 35). $\blacksquare$

DEFINITION 3 **Orthogonal Vectors** The nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ are called **orthogonal** (or **perpendicular**) if the angle between them is $\pi/2$.

EXAMPLE 3 **Two Orthogonal Vectors** Show that the vectors $\mathbf{u} = 3\mathbf{i} - 4\mathbf{j}$ and $\mathbf{v} = 4\mathbf{i} + 3\mathbf{j}$ are orthogonal.

Solution $\mathbf{u} \cdot \mathbf{v} = 3 \cdot 4 - 4 \cdot 3 = 0$. This implies that $\cos \varphi = (\mathbf{u} \cdot \mathbf{v})/(|\mathbf{u}|\,|\mathbf{v}|) = 0$. Since φ is in the interval $[0, \pi]$, $\varphi = \pi/2$. $\blacksquare$

THEOREM 4 The nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.

Proof This proof is also left as an exercise (see Problem 36). $\blacksquare$

---

† These numbers, like others in the text, were obtained with a hand calculator.

‡ When doing this computation yourself, make certain that your calculator is set to radian mode.

EXAMPLE 4 **Determining a Way to Make Two Vectors Orthogonal or Parallel**
Let $\mathbf{u} = \mathbf{i} + 4\mathbf{j}$ and $\mathbf{v} = 3\mathbf{i} + \alpha\mathbf{j}$. Determine α such that (i) $\mathbf{u}$ and $\mathbf{v}$ are orthogonal; (ii) $\mathbf{u}$ and $\mathbf{v}$ are parallel.

Solution

i. We have $\mathbf{u} \cdot \mathbf{v} = 3 + 4\alpha$. For $\mathbf{u}$ and $\mathbf{v}$ to be orthogonal, we must have $\mathbf{u} \cdot \mathbf{v} = 0$. This implies that $3 + 4\alpha = 0$ or $\alpha = -\frac{3}{4}$.

ii. Here we must have $\varphi = 0$ or π so that $\cos\varphi = \pm 1$. Then

$$\cos\varphi = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{3 + 4\alpha}{\sqrt{17}\sqrt{9 + \alpha^2}} = \pm 1$$

Squaring both sides of this last equation, we obtain

$$9 + 24\alpha + 16\alpha^2 = 17(9 + \alpha^2) = 153 + 17\alpha^2.$$

This leads to the quadratic equation

$$\alpha^2 - 24\alpha + 144 = 0 = (\alpha - 12)^2,$$

with the single solution $\alpha = 12$.

Check. $3\mathbf{i} + 12\mathbf{j} = 3(\mathbf{i} + 4\mathbf{j})$, so that $\mathbf{v} = 3\mathbf{u}$ and $\mathbf{u}$ and $\mathbf{v}$ are parallel by Theorem 3. ■

A number of interesting problems involve the notion of the projection of one vector along another. Before defining this, we prove the following theorem.

THEOREM 5 Let $\mathbf{v}$ be a nonzero vector. Then for any other vector $\mathbf{u}$, the vector

$$\mathbf{w} = \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}\mathbf{v}$$

is orthogonal to $\mathbf{v}$.

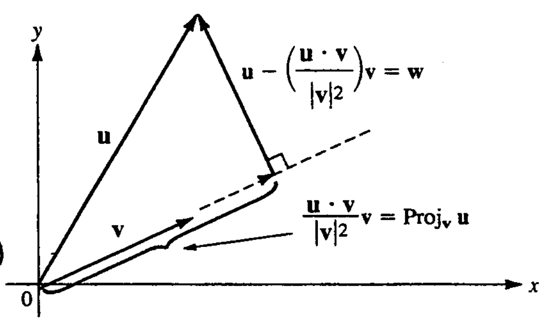
Proof

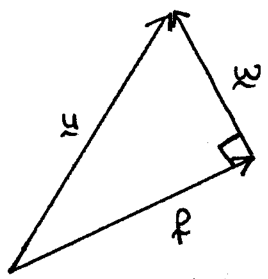
$$\mathbf{w} \cdot \mathbf{v} = \left[\mathbf{u} - \frac{(\mathbf{u} \cdot \mathbf{v})\mathbf{v}}{|\mathbf{v}|^2}\right] \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{v} - \frac{(\mathbf{u} \cdot \mathbf{v})(\mathbf{v} \cdot \mathbf{v})}{|\mathbf{v}|^2}$$

$$= \mathbf{u} \cdot \mathbf{v} - \frac{(\mathbf{u} \cdot \mathbf{v})|\mathbf{v}|^2}{|\mathbf{v}|^2} = \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{v} = 0$$ ■

The vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are illustrated in Figure 3.14.

Given $\underset{\sim}{u}$ and $\underset{\sim}{v}$, suppose we want to write $\underset{\sim}{u}$ as a sum, $\underset{\sim}{u} = \underset{\sim}{p} + \underset{\sim}{w}$, where $\underset{\sim}{p}$ and $\underset{\sim}{w}$ are perpendicular (orthogonal) as shown here.





$\underset{\sim}{u} = \underset{\sim}{p} + \underset{\sim}{w}$

Figure 3.14 The vector $\mathbf{w} = \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}\mathbf{v}$ is orthogonal to $\mathbf{v}$

If we know $\underset{\sim}{p}$, then $\underset{\sim}{w}$ is simply $\underset{\sim}{u} - \underset{\sim}{p}$.
This Theorem 5 proves that $\underset{\sim}{p}$ should be $\frac{u \cdot v}{|v|^2}\underset{\sim}{v}$, so that Defn. 4 follows.

DEFINITION 4 **Projection** Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors. Then the **projection of u on v** is a vector, denoted by $\text{proj}_{\mathbf{v}}\, \mathbf{u}$, which is defined by

(See Study Guide, pp 12-13, for a more detailed development of these 2 formulae, without Theorem 5.)

$$\boxed{\text{proj}_{\mathbf{v}}\, \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}\mathbf{v}} \tag{4}$$

$$\boxed{\text{The component of } \mathbf{u} \text{ in the direction } \mathbf{v} \text{ is } \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|}} \tag{5}$$

Note that $\mathbf{v}/|\mathbf{v}|$ is a unit vector in the direction of $\mathbf{v}$.

Remark 1. From Figures 3.14 and 3.15 and the fact that $\cos \varphi = (\mathbf{u} \cdot \mathbf{v})/(|\mathbf{u}|\ |\mathbf{v}|)$, we find that

> $\mathbf{v}$ and $\text{proj}_{\mathbf{v}}\, \mathbf{u}$ have (*i*) the same direction if $\mathbf{u} \cdot \mathbf{v} > 0$ and (*ii*) opposite directions if $\mathbf{u} \cdot \mathbf{v} < 0$.

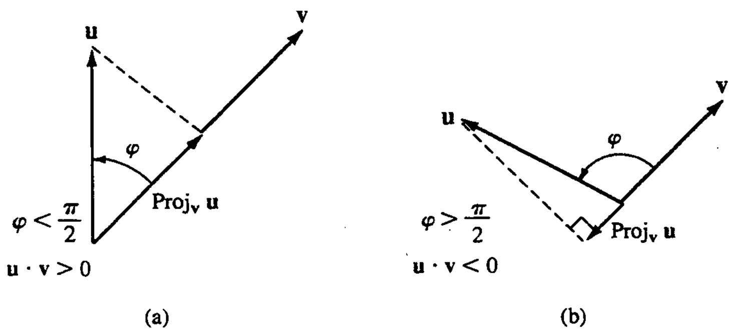


Figure 3.15 (a) $\mathbf{v}$ and $\text{Proj}_{\mathbf{v}}\, \mathbf{u}$ have the same direction if $\mathbf{u} \cdot \mathbf{v} > 0$. (b) $\mathbf{v}$ and $\text{proj}_{\mathbf{v}}\, \mathbf{u}$ have opposite directions if $\mathbf{u} \cdot \mathbf{v} < 0$

Remark 2. $\text{Proj}_{\mathbf{v}}\, \mathbf{u}$ can be thought of as the "$\mathbf{v}$-component" of the vector $\mathbf{u}$.

Remark 3. If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, then $\mathbf{u} \cdot \mathbf{v} = 0$ so that $\text{proj}_{\mathbf{v}}\, \mathbf{u} = \mathbf{0}$.

Remark 4. An alternative definition of projection is: If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, then $\text{proj}_{\mathbf{v}}\, \mathbf{u}$ is the unique vector having these properties:

vector p — **i.** $\text{Proj}_{\mathbf{v}}\, \mathbf{u}$ is parallel to $\mathbf{v}$.
vector w — **ii.** $\mathbf{u} - \text{proj}_{\mathbf{v}}\, \mathbf{u}$ is orthogonal to $\mathbf{v}$.

We say that u is "decomposed" into the sum of vectors p and w.