Differential vector calculus

P. R. Baxandall &
H. Liebeck

Longman Mathematical Texts

Differential vector calculus

P.R. Baxandall and H. Liebeck

Department of Mathematics,
The University of Keele

Longman

London and New York

Longman Group Limited
Longman House
Burnt Mill, Harlow, Essex, UK

Published in the United States of America
by Longman Inc., New York

First published 1981

British Library Cataloguing in Publication Data

Differential vector calculus.—(Longman
mathematical texts).
1. Calculus, Differential
2. Vector analysis
I. Baxandall, P
II. Liebeck, Hans
515'.63 QA304 80-41123
ISBN 0-582-44193-5

Set in Monophoto Times New Roman
Printed in Great Britain by
Butler & Tanner Ltd, Frome and London

Preface

This book provides an introduction to advanced differential calculus set against a background of linear algebra. It is designed as a first or second year course for undergraduates who have some knowledge of linear algebra and real analysis.

Traditionally linear algebra, vector analysis and calculus of functions of several variables are taught as separate subjects. They are, however, closely related. The underlying links are established by defining the differentiability of a vector function at a point in its domain in terms of the existence of an approximating linear transformation (differential). Many of the important classical results of vector calculus are essentially concerned with the geometrical properties of this differential and of its image and graph.
The Chain Rule for example, which plays an important part in the development of the theory, is merely an expression of the fact that the linear approximation to the composition of two functions is the composition of their individual linear approximations. This theorem enables us to study the relationship between different parametrizations of curves and surfaces.

In the study of vector functions from $\mathbb{R}^m$ to $\mathbb{R}^n$ the cases $m = 1$ and $n = 1$ warrant separate discussion. After a short introductory chapter we proceed in Chapter 2 to consider functions from $\mathbb{R}$ to $\mathbb{R}^n$ (including the study of curves, differential geometry and dynamics). Chapter 3 deals with functions from $\mathbb{R}^m$ to $\mathbb{R}$ (real-valued functions of many variables, Taylor's Theorem and applications). Finally, the general theory of functions from $\mathbb{R}^m$ to $\mathbb{R}^n$ is covered in Chapter 4. In particular we prove the important inverse function and implicit function theorems. Our approach has the advantage of introducing the concept of differentiability of vector functions in easy stages. Certain theorems (the Chain Rule in particular) appear at a number of points in progressively more general form.

We have tried to give readable yet rigorous proofs (often omitted from introductory texts) of some of the important classical theorems. The reader is recommended to attempt as many of the exercises as he can. Apart from the usual routine applications of definitions and theorems, many of the exercises explain by way of counter-example the significance of the hypotheses of the theorems.

We wish to express our thanks to Professor Alan Jeffrey editor of Longman Mathematical Texts for inviting us to contribute to the series. We are grateful to many colleagues and friends for fruitful discussions concerning the text and in particular to the son of one of the authors, Dr Martin Liebeck, who kindly read the whole manuscript and suggested many improvements. Finally, very special thanks are due to Miss Christine Williams for typing the manuscript, and for cheerfully retyping parts of it as the authors settled their differences.

Keele
March 1980

P.R. Baxandall
H. Liebeck

Longman Mathematical Texts

Edited by Alan Jeffrey

Elementary mathematical analysis I.T. Adamson
An introduction to the theory of elasticity R.J. Atkin and N. Fox
Differential vector calculus P.R. Baxandall and H. Liebeck
Set theory and abstract algebra T.S. Blyth
The theory of ordinary differential equations J.C. Burkill
Random variables L.E. Clarke
Electricity C.A. Coulson and T.J.M. Boyd
Waves C.A. Coulson and A. Jeffrey
A short course in general relativity J. Foster and J D Nightingale
Optimisation D.M Greig
Integral equations B.L. Moiseiwitsch
Functions of a complex variable E.G. Phillips
Special functions of mathematical physics and chemistry I.N. Sneddon
Continuum mechanics A.J.M. Spencer

Contents

1

Basic linear algebra and analysis

1.1 Introduction

Basic analysis and linear algebra play a fundamental role in the generalization of elementary calculus to the theory of vector-valued functions. We shall assume familiarity with the analysis and linear algebra usually found in a first course. Most of this background material is summarized in this chapter.

A first course in calculus deals with real-valued functions of one real variable. Such a function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ is defined on a domain D which is a subset of the real line $\mathbb{R}$. The value that f takes at $x \in D$ is a real number denoted by $f(x) \in \mathbb{R}$. For example, the rule

$$f(x) = \sqrt{(1 - x^2)}$$

can be taken to define a real-valued function $f : D \subseteq \mathbb{R} \to \mathbb{R}$, where D is the closed interval $[-1, 1] = \{x \in \mathbb{R} | -1 \leqslant x \leqslant 1\}$.

Let $\mathbb{R}^m$ denote the set of all m-tuples of real numbers $(x_1, x_2, \ldots, x_m)$, $x_i \in \mathbb{R}$, $i = 1, \ldots, m$. (In particular, $\mathbb{R}^1 = \mathbb{R}$.) We shall be concerned with the study of functions $f : D \subseteq \mathbb{R}^m \to \mathbb{R}^n$ defined on a subset D of $\mathbb{R}^m$ and taking values in $\mathbb{R}^n$. For example, the rule

1.1.1 $$f(x_1, x_2, x_3) = (\sqrt{(1 - x_3^2)}, x_1 x_2 x_3)$$

can be taken to define a function $f : D \subseteq \mathbb{R}^3 \to \mathbb{R}^2$, where D is the subset of $\mathbb{R}^3$ given by

1.1.2 $$D = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | -1 \leqslant x_3 \leqslant 1\}.$$

The set $\mathbb{R}^m$ defined above is given the structure of a real vector space by defining an addition of m-tuples

1.1.3 $$(x_1, \ldots, x_m) + (y_1, \ldots, y_m) = (x_1 + y_1, \ldots, x_m + y_m)$$

and a multiplication by scalars in $\mathbb{R}$,

1.1.4 $$k(x_1, \ldots, x_m) = (kx_1, \ldots, kx_m), \qquad k \in \mathbb{R}.$$

Viewed in this way, the rule 1.1.1 defines a funtion f on a subset D

of the vector space $\mathbb{R}^3$ such that the values that f takes lie in the vector space $\mathbb{R}^2$.

With $\mathbb{R}^m$ interpreted as a vector space, the study of functions $f: D \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is appropriately called vector calculus. (The alternative title 'calculus of functions of several variables' is sometimes preferred when $m > 1$.) We observe that vector calculus includes the case $m = n = 1$ and so is a generalization of elementary calculus.

Exercise 1.1

1. Suggest a possible subset $D \subseteq \mathbb{R}^3$ as the domain of a function $f: D \subseteq \mathbb{R}^3 \to \mathbb{R}^2$ which is given by the rule
   (a) $f(x_1, x_2, x_3) = (x_1/x_2, \sqrt{(1 - x_3^2)})$;
   (b) $f(x, y, z) = (e^{1/x}\tan(xyz), x^2 + y^2 + z^2)$;
   (c) $f(x, y, z) = (1/(x^2 - z^2), \ln(xyz))$.

   Answers:
   (a) $\{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_2 \neq 0, x_3 \in [-1, 1]\}$;
   (b) $\{(x, y, z) \in \mathbb{R}^3 | x \neq 0, xyz \neq (k + \frac{1}{2})\pi$, for integers $k\}$;
   (c) $\{(x, y, z) \in \mathbb{R}^3 | x \neq \pm z, xyz > 0\}$.

1.2 The vector space $\mathbb{R}^m$

In section 1.1 we pointed out that the set $\mathbb{R}^m$ of all m-tuples $(x_1, \ldots, x_m)$, $x_i \in \mathbb{R}$, $i = 1, \ldots, m$, can be regarded as a vector space over $\mathbb{R}$ if we impose the rules 1.1.3, 1.1.4 of addition and scalar multiplication. We shall often denote a vector of $\mathbb{R}^m$ by a single letter in bold-face, thus: $\mathbf{x} = (x_1, \ldots, x_m)$. In $\mathbb{R}^2$ and $\mathbb{R}^3$ the familiar notation $\mathbf{r} = (x, y)$ and $\mathbf{r} = (x, y, z)$ is useful.

Consider the vectors in $\mathbb{R}^m$

$$\mathbf{e}_1 = (1, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, \ldots, 0), \ldots, \mathbf{e}_m = (0, 0, \ldots, 1),$$

where $\mathbf{e}_i$ has 1 in the ith place and 0 elsewhere. The set $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ is clearly a basis of the vector space $\mathbb{R}^m$, since any $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$ has the unique expression $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m$ as a linear combination of the $\mathbf{e}_i$'s. In view of its simple form we call the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ the *standard* (or *natural*) *basis* of $\mathbb{R}^m$. No other bases of $\mathbb{R}^m$ are used in this book.

The vector spaces $\mathbb{R}^1 = \mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}^3$ are conveniently pictured as a number line, as a plane and as three-dimensional space. For example, we picture $\mathbb{R}^2$ by choosing an ordered pair of perpendicular axes and a unit of length, and associating the vector $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$

with the point in the plane whose coordinates relative to the axes are (x_1, x_2) in the usual way.

An important alternative way of picturing the vector $\mathbf{x} = (x_1, x_2)$ in the plane is by an arrow joining the origin (0, 0) to the point labelled (x_1, x_2). Both ways of picturing vectors will frequently be used – sometimes in the same diagram. The arrow representation is particularly important in physical applications, for example when we wish to picture velocities, accelerations or forces.

In the arrow representation, the rule 1.1.3 of vector addition is the well known parallelogram law of vector addition. See Fig. 1.1(i). The arrows joining O to P and Q to R are identical in all respects except for their position in the plane. We therefore agree to picture $\mathbf{x} = (x_1, x_2)$ not only by the arrow $\overrightarrow{OP}$ but also by an arrow joining (y_1, y_2) to $(x_1 + y_1, x_2 + y_2)$, where y_1, y_2 are arbitrarily chosen real numbers.

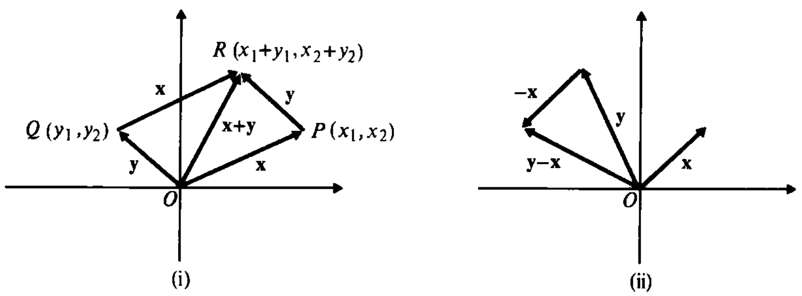

Fig. 1.1

Vector subtraction is illustrated in Fig. 1.1(ii). We have, by definition, $\mathbf{y} - \mathbf{x} = \mathbf{y} + (-\mathbf{x})$, where $-\mathbf{x} = (-x_1, -x_2)$.

Considerations similar to the above apply to a pictorial representation of $\mathbb{R}^3$ relative to three mutually perpendicular axes.

1.2.1 ***Definition.*** *Given vectors* $\mathbf{x} = (x_1, \ldots, x_m)$ *and* $\mathbf{y} = (y_1, \ldots, y_m)$ *in* $\mathbb{R}^m$, *the* dot product (*or* scalar product) *of* $\mathbf{x}$ *and* $\mathbf{y}$ *is defined to be the real number*

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_m y_m.$$

The dot product has the following important properties

1.2.2 *Symmetry*: $\mathbf{x}\cdot\mathbf{y}=\mathbf{y}\cdot\mathbf{x}$, $\quad \mathbf{x}, \mathbf{y}\in\mathbb{R}^m$

1.2.3 *Linearity*: $(k\mathbf{x}+l\mathbf{y})\cdot\mathbf{z}=k(\mathbf{x}\cdot\mathbf{z})+l(\mathbf{y}\cdot\mathbf{z})$,
$\mathbf{x}, \mathbf{y}, \mathbf{z}\in\mathbb{R}^m$, $k, l\in\mathbb{R}$

1.2.4 *Positivity*: $\mathbf{x}\cdot\mathbf{x}>0$, for all $\mathbf{x}\neq\mathbf{0}, \mathbf{x}\in\mathbb{R}^m$.

1.2.5 *Definition.* *The* norm *or* length of $\mathbf{x}\in\mathbb{R}^m$ *is the non-negative real number* $\|\mathbf{x}\|\geqslant 0$ *such that*

$$\|\mathbf{x}\|^2=\mathbf{x}\cdot\mathbf{x}=x_1^2+\ldots+x_m^2.$$

1.2.6 *Example.* If $\mathbf{a}=(2, 1)$, then $\|\mathbf{a}\|^2=2^2+1^2=5$ and $\|\mathbf{a}\|=\sqrt{5}$.
Note that $\|\mathbf{x}\|=0$ if and only if $\mathbf{x}=\mathbf{0}$.

1.2.7 *Example.* It follows from Definition 1.2.5 that for $\mathbf{x}\in\mathbb{R}^m$, $k\in\mathbb{R}$, $\|k\mathbf{x}\|=|k|\,\|\mathbf{x}\|$. For example, $\|-2\mathbf{x}\|=2\,\|\mathbf{x}\|$.

The following theorem will be used extensively in this book.

1.2.8 *Theorem.* The Cauchy–Schwarz inequality. *For any vectors* $\mathbf{x}, \mathbf{y}\in\mathbb{R}^m$

1.2.9 $$|\mathbf{x}\cdot\mathbf{y}|\leqslant\|\mathbf{x}\|\,\|\mathbf{y}\|,$$

with equality if and only if the vectors $\mathbf{x}, \mathbf{y}$ *are linearly dependent.*

Proof. [*i*] If $\mathbf{y}=\mathbf{0}$, then $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent, and 1.2.9 is an equality, both sides being zero.

[*ii*] If $\mathbf{y}\neq\mathbf{0}$ and $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent, then there exists $k\in\mathbb{R}$ such that $\mathbf{x}=k\mathbf{y}$. In this case 1.2.9 is an equality, both sides being equal to $|k|\,\|\mathbf{y}\|^2$.

[*iii*] The remaining case is that $\mathbf{x}$ and $\mathbf{y}$ are linearly independent. Then, for any $k\in\mathbb{R}$, $\mathbf{x}+k\mathbf{y}\neq\mathbf{0}$ and, by the properties 1.2.2–1.2.4 of the dot product,

$$\begin{aligned}0<\|\mathbf{x}-k\mathbf{y}\|^2&=(\mathbf{x}+k\mathbf{y})\cdot(\mathbf{x}+k\mathbf{y})\\&=\|\mathbf{x}\|^2+2k(\mathbf{x}\cdot\mathbf{y})+k^2\|\mathbf{y}\|^2.\end{aligned}$$

With $k=-(\mathbf{x}\cdot\mathbf{y})/\|\mathbf{y}\|^2$, a simple calculation results in 1.2.9 as a strict inequality.

The Cauchy–Schwarz inequality is often stated in the form

1.2.10 $$\left(\sum_{i=1}^{m}x_iy_i\right)^2\leq\sum_{i=1}^{m}x_i^2\sum_{i=1}^{m}y_i^2.$$

This follows by squaring 1.2.9 and applying the definitions of dot product and norm.

1.2.11 ***Theorem.*** The Triangle Inequality. *For any vectors* **x**, **y** *in* $\mathbb{R}^m$,

$$\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|.$$

Proof

1.2.12
$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) = \|\mathbf{x}\|^2 + 2\mathbf{x}\cdot\mathbf{y} + \|\mathbf{y}\|^2 \\ &\leqslant \|\mathbf{x}\|^2 + 2|\mathbf{x}\cdot\mathbf{y}| + \|\mathbf{y}\|^2 \\ &\leqslant \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2 \qquad \text{by (1.2.9)} \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.\end{aligned}$$

Taking square roots on both sides, we obtain the Triangle Inequality.

A picture for the cases $\mathbb{R}^2$ and $\mathbb{R}^3$ shows why Theorem 1.2.11 is called the Triangle Inequality, for it is related to the property of triangles that the sum of the lengths $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$ of two sides of a triangle is not smaller than the length $\|\mathbf{x}+\mathbf{y}\|$ of the third side. (See Fig. 1.1.)

1.2.13 ***Corollary.*** [i] *For any* **x** *and* **y** *in* $\mathbb{R}^m$

1.2.14
$$\big|\,\|\mathbf{x}\| - \|\mathbf{y}\|\,\big| \leqslant \|\mathbf{x}-\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|.$$

[ii] *For any n vectors* $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ *and scalars* $x_1, \dots, x_n \in \mathbb{R}$,

1.2.15
$$\|x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n\| \leqslant |x_1|\,\|\mathbf{a}_1\| + \cdots + |x_n|\,\|\mathbf{a}_n\|.$$

[iii] *For any* $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^m$,

1.2.16
$$\|(x_1, \dots, x_m)\| \leqslant |x_1| + \cdots + |x_m|.$$

Proof. Exercise.

The well-known cosine rule applied to the triangle OPR of Fig. 1.1 gives

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta,$$

where θ is the angle between the non-zero vectors **x** and **y**. Comparing this expression with 1.2.12 we obtain

$$\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

In particular, the vector $\mathbf{x}$ is perpendicular (or orthogonal) to $\mathbf{y}$ if and only if $\mathbf{x}\cdot\mathbf{y}=0$. The following generalization applies to $\mathbb{R}^m$.

1.2.17 *Definition.* *The vector* $\mathbf{x}\in\mathbb{R}^m$ *is* orthogonal *to the vector* $\mathbf{y}\in\mathbb{R}^m$ *if and only if* $\mathbf{x}\cdot\mathbf{y}=0$.

Note that by the symmetry of the dot product, $\mathbf{x}$ is orthogonal to $\mathbf{y}$ if and only if $\mathbf{y}$ is orthogonal to $\mathbf{x}$.

1.2.18 *Example.* *Equation of a plane in* $\mathbb{R}^3$. A plane in $\mathbb{R}^3$ is specified by a point $\mathbf{q}\in\mathbb{R}^3$ in the plane and a non-zero vector $\mathbf{n}\in\mathbb{R}^3$ perpendicular (or *normal*) to the plane.

The point $\mathbf{r}=(x, y, z)\in\mathbb{R}^3$ lies in the plane if and only if $\mathbf{r}-\mathbf{q}$ is orthogonal to $\mathbf{n}$, that is, if and only if

1.2.19 $$(\mathbf{r}-\mathbf{q})\cdot\mathbf{n}=0.$$

Equation 1.2.19 is called the equation of the plane containing the point $\mathbf{q}$ and having normal $\mathbf{n}$.

For example, the equation of the plane containing $\mathbf{q}=(1, 1, 1)$ with normal $\mathbf{n}=(2, 4, 6)$ is $(x, y, z)\cdot(2, 4, 6)=(1, 1, 1)\cdot(2, 4, 6)$, that is,

$$2x+4y+6z=12.$$

The dot product is defined on vectors in $\mathbb{R}^m$, where m is arbitrary. The following vector product is defined in $\mathbb{R}^3$ only.

1.2.20 *Definition.* *The* vector product *of* $\mathbf{b}=(b_1, b_2, b_3)$ *and* $\mathbf{c}=(c_1, c_2, c_3)$ *in* $\mathbb{R}^3$ *is the vector*

$$\mathbf{b}\wedge\mathbf{c}=(b_2c_3-b_3c_2, b_3c_1-b_1c_3, b_1c_2-b_2c_1).$$

The formula for $\mathbf{b}\wedge\mathbf{c}$ is conveniently obtained by expanding the formal determinant

1.2.21 $$\mathbf{b}\wedge\mathbf{c}=\begin{vmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix},$$

where $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is the standard basis of $\mathbb{R}^3$. An alternative common notation for the standard basis of $\mathbb{R}^3$ is $\mathbf{i}=(1, 0, 0)$, $\mathbf{j}=(0, 1, 0)$, $\mathbf{k}=(0, 0, 1)$ and we shall use it occasionally.

We state the following standard results about the vector product without proof.

1.2.22 *Theorem.* *Let* $\mathbf{b}$ *and* $\mathbf{c}$ *be vectors in* $\mathbb{R}^3$. *Then*

[i] *the vector* $\mathbf{b}\wedge\mathbf{c}$ *is orthogonal to* $\mathbf{b}$ *and to* $\mathbf{c}$;

[ii] **b** *and* **c** *are linearly dependent if and only if* $\mathbf{b} \wedge \mathbf{c} = \mathbf{0}$;

[iii] *if* **b** and **c** *are linearly independent then relative to a right-handed coordinate system of* $\mathbb{R}^3$ *the vectors* **b**,**c**, $\mathbf{b} \wedge \mathbf{c}$ *from a right-handed triple of vectors*;

[iv] $\|\mathbf{b} \wedge \mathbf{c}\| = \|\mathbf{b}\| \, \|\mathbf{c}\| \sin \phi$,

where ϕ *is the angle between* **b** *and* **c**. *Thus* $\|\mathbf{b} \wedge \mathbf{c}\|$ *measures the area of the parallelogram with* **b** *and* **c** *as adjacent sides.*

Exercises 1.2

1. Prove Corollary 1.2.13. (Hint: to prove the right-hand inequality 1.2.14 apply Theorem 1.2.11 to $\mathbf{x} + (-\mathbf{y})$; for the left-hand inequality put $\mathbf{x} - \mathbf{y} = \mathbf{z}$.)

2. (a) Prove that the zero vector in $\mathbb{R}^m$ is orthogonal to every vector $\mathbf{x} \in \mathbb{R}^m$.
   (b) Prove that if $\mathbf{x} \in \mathbb{R}^m$ is orthogonal to the vectors $\mathbf{y}_1, \ldots, \mathbf{y}_r$ in $\mathbb{R}^m$, then **x** is orthogonal to every vector in the subspace of $\mathbb{R}^m$ spanned by $\mathbf{y}_1, \ldots, \mathbf{y}_r$.

3. Find the angle between (a) the vectors $(1, -1, 0)$ and $(-4, 1, 1)$; (b) the vectors $(1, -1, 0)$ and $(4, -1, -1)$. Compare.

Answer: (a) $\cos \theta = -\frac{5}{6}$; (b) $\cos \phi = \frac{5}{6}$. $\theta + \phi = \pi$.

4. Prove that vectors $\mathbf{x} - \mathbf{y}$ and $\mathbf{x} + \mathbf{y}$ in $\mathbb{R}^m$ are orthogonal if and only if $\|\mathbf{x}\| = \|\mathbf{y}\|$. Illustrate this result in $\mathbb{R}^2$, and deduce that the diagonals of a parallelogram intersect at right angles if and only if the parallelogram is a rhombus.

5. Find the equation of the plane in $\mathbb{R}^3$
   (a) containing the point $(1, -1, 1)$ and with normal $\mathbf{n} = (0, 1, 1)$;
   (b) containing the points $(1, 0, 1)$, $(0, 1, 1)$ and $(1, 1, 0)$.

Answers: (a) $y + z = -2$; (b) $x + y + z = 2$.

6. Prove from the definition of the vector product that
   (a) $\mathbf{b} \wedge \mathbf{c} = -\mathbf{c} \wedge \mathbf{b}$; (b) $\mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

7. Find a vector orthogonal to both $\mathbf{a} = (1, -1, 2)$ and $\mathbf{b} = (2, 0, 1)$.

Answer: any scalar multiple of the vector product $\mathbf{a} \wedge \mathbf{b} = (-1, 3, 2)$.

1.3 Linear functions

1.3.1 ***Definition.*** *A* linear *function* $L : \mathbb{R}^m \to \mathbb{R}^n$ *with domain* $\mathbb{R}^m$ *and codomain* $\mathbb{R}^n$ *is a rule that assigns to each* $\mathbf{x} \in \mathbb{R}^m$ *a unique vector* $L(\mathbf{x}) \in \mathbb{R}^n$ *such that for any* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ *and* $k, l \in \mathbb{R}$,

1.3.2 $$L(k\mathbf{x} + l\mathbf{y}) = kL(\mathbf{x}) + lL(\mathbf{y}).$$

It follows from 1.3.2 (by induction on r) that for any $\mathbf{a}_1, \ldots, \mathbf{a}_r \in \mathbb{R}^m$ and $k_1, \ldots, k_r \in \mathbb{R}$,

1.3.3 $$L(k_1\mathbf{a}_1 + \cdots + k_r\mathbf{a}_r) = k_1 L(\mathbf{a}_1) + \cdots + k_r L(\mathbf{a}_r).$$

1.3.4 *Theorem.* *A linear function $L : \mathbb{R}^m \to \mathbb{R}^n$ is completely determined by its effect on the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$ of $\mathbb{R}^m$. Moreover, an arbitrary choice of vectors $L(\mathbf{e}_1), \ldots, L(\mathbf{e}_m)$ in $\mathbb{R}^n$ determines a linear function $L : \mathbb{R}^m \to \mathbb{R}^n$.*

Proof. Choose $L(\mathbf{e}_i) \in \mathbb{R}^n$ for each $i = 1, \ldots, n$. Then for $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$,

$$\begin{aligned} L(\mathbf{x}) &= L(x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m) \\ &= x_1 L(\mathbf{e}_1) + \cdots + x_m L(\mathbf{e}_m), \qquad \text{by (1.3.3).} \end{aligned}$$

Therefore the image under L of any $\mathbf{x} \in \mathbb{R}^m$ is known, and so L is completely determined.

Note that a linear function has a vector space $\mathbb{R}^m$ as its domain, whereas a non-linear function may be defined on a subset of $\mathbb{R}^m$. See, for example, 1.1.1.

The *image* of a function $L : \mathbb{R}^m \to \mathbb{R}^n$ is defined as the set of all image vectors:

1.3.5 $$\operatorname{im} L = \{L(\mathbf{x}) \in \mathbb{R}^n \mid \mathbf{x} \in \mathbb{R}^m\}.$$

We say that L maps $\mathbb{R}^m$ *onto* $\mathbb{R}^n$ if $\operatorname{im} L = \mathbb{R}^n$.

The *kernel* of $L : \mathbb{R}^m \to \mathbb{R}^n$ is defined as the set of all $\mathbf{x} \in \mathbb{R}^m$ that are mapped by L to zero:

1.3.6 $$\ker L = \{\mathbf{x} \in \mathbb{R}^m \mid L(\mathbf{x}) = \mathbf{0} \in \mathbb{R}^n\}.$$

When $L : \mathbb{R}^m \to \mathbb{R}^n$ is linear, then $\operatorname{im} L$ is a subspace of $\mathbb{R}^n$, $\ker L$ is a subspace of $\mathbb{R}^m$, and the dimensions of $\operatorname{im} L$ and $\ker L$ are related by the celebrated formula (which we leave unproved)

1.3.7 $$\dim \operatorname{im} L + \dim \ker L = \dim \mathbb{R}^m = m.$$

1.3.8 *Definition.* *A function f defined on a domain D is said to be* 1–1 (one-to-one) *on D if distinct elements of D have distinct images under f; that is if, for any $\mathbf{x} \in D$, $\mathbf{y} \in D$, $\mathbf{x} \neq \mathbf{y}$ implies that $f(\mathbf{x}) \neq f(\mathbf{y})$.*

1.3.9 *Theorem.* *A linear function $L : \mathbb{R}^m \to \mathbb{R}^n$ is 1–1 if and only if* $\ker L = \{\mathbf{0}\}$.

Proof. Exercise.

We shall require (in section 4.6) the following definition and result concerning a linear function $L : \mathbb{R}^n \to \mathbb{R}^n$ whose domain and co-domain are the same space $\mathbb{R}^n$.

1.3.10 *Definition.* *An* isomorphism *on* $\mathbb{R}^n$ *is a linear function* $L : \mathbb{R}^n \to \mathbb{R}^n$ *mapping* $\mathbb{R}^n$ *onto itself.*

1.3.11 *Theorem.* *A linear function* $L : \mathbb{R}^n \to \mathbb{R}^n$ *is an isomorphism on* $\mathbb{R}^n$ *if and only if* L *is* 1–1.

Proof. By Definition 1.3.10, L is an isomorphism on $\mathbb{R}^n$ if and only if $\operatorname{im} L = \mathbb{R}^n$. By 1.3.7 (applied to $\mathbb{R}^n$ as domain) this is so if and only if $\ker L = \{\mathbf{0}\}$. The theorem follows from Theorem 1.3.9.

We now outline the procedure for representing a linear function $L : \mathbb{R}^m \to \mathbb{R}^n$ by a matrix. By Theorem 1.3.4, the function L is determined by the images under L of the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$ of its domain $\mathbb{R}^m$. We must avoid confusing the standard bases of $\mathbb{R}^m$ and of $\mathbb{R}^n$, so let us denote the standard basis of $\mathbb{R}^n$ by $\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*$. Suppose that for each $j = 1, \ldots, m$,

$$L(\mathbf{e}_j) = a_{1j}\mathbf{e}_1^* + \cdots + a_{nj}\mathbf{e}_n^*, \qquad a_{ij} \in \mathbb{R}, \quad i = 1, \ldots, n.$$

Then for any $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$, if $L(\mathbf{x}) = \mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$,

1.3.12
$$y_i = \sum_{j=1}^{m} a_{ij} x_j, \qquad i = 1, \ldots, n.$$

Formula 1.3.12 is conveniently written in matrix form

1.3.13
$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix},$$

the evaluation being performed by the usual matrix multiplication.

We say that a vector $\mathbf{z} = (z_1, \ldots, z_q) \in \mathbb{R}^q$ is represented relative to the standard basis of $\mathbb{R}^q$ by the column ($q \times 1$ matrix)

$$[\mathbf{z}] = \begin{bmatrix} z_1 \\ \vdots \\ z_q \end{bmatrix}$$

and that the linear function $L : \mathbb{R}^m \to \mathbb{R}^n$ discussed above is