10213

10214

10215

10216

10217

FEUE WHEP
10218

Useful Formulae

$$\hbar = 6.58 \times 10^{-25} \text{ GeV sec} = 1$$
$$\hbar c = 0.197 \text{ GeV F} = 1$$
$$(1 \text{ GeV})^{-2} = 0.389 \text{ mb}$$
$$\alpha = \frac{e^2}{4\pi} = \frac{1}{137}$$

$$x^\mu = (t, \mathbf{x}), \qquad p^\mu = (E, \mathbf{p}) = i\left(\frac{\partial}{\partial t}, -\nabla\right) = i\partial^\mu$$

$$p \cdot x = Et - \mathbf{p}\cdot\mathbf{x}, \qquad p^2 \equiv p^\mu p_\mu = E^2 - \mathbf{p}^2 = m^2$$

$$(\Box^2 + m^2)\phi = 0, \qquad (i\gamma^\mu \partial_\mu - m)\psi = 0.$$

In an electromagnetic field, $i\partial^\mu \to i\partial^\mu + eA^\mu$ (charge $-e$)

$$j^\mu = -ie(\phi^* \partial^\mu \phi - \phi\, \partial^\mu \phi^*), \qquad j^\mu = -e\bar{\psi}\gamma^\mu\psi$$

γ-Matrices

$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}, \qquad \gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0.$

$\gamma^{0\dagger} = \gamma^0, \qquad \gamma^0\gamma^0 = I; \qquad \gamma^{k\dagger} = -\gamma^k, \qquad \gamma^k\gamma^k = -I, \qquad k = 1,2,3.$

$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3, \qquad \gamma^\mu\gamma^5 + \gamma^5\gamma^\mu = 0, \qquad \gamma^{5\dagger} = \gamma^5.$

(Trace theorems on pages 123 and 261)

Standard representation:

$$\gamma^0 \equiv \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \qquad \boldsymbol{\gamma} \equiv \beta\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}, \qquad \gamma^5 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Spinors

$$\begin{matrix} (\not{p} - m)u = 0 \\ \bar{u}(\not{p} - m) = 0 \end{matrix} \qquad \begin{cases} \bar{u} \equiv u^\dagger\gamma^0 \\ \not{p} \equiv \gamma_\mu p^\mu \end{cases}$$

$$u^{(r)\dagger}u^{(s)} = 2E\delta_{rs}, \qquad \bar{u}^{(r)}u^{(s)} = 2m\delta_{rs}, \qquad \sum_{s=1,2} u^{(s)}\bar{u}^{(s)} = \not{p} + m = 2m\Lambda_+$$

$\frac{1}{2}(1 - \gamma^5)u \equiv u_L, \qquad \frac{1}{2}(1 + \gamma^5)u \equiv u_R.$

If $m = 0$ or $E \gg m$, then u_L has helicity $\lambda = -\frac{1}{2}$, u_R has $\lambda = +\frac{1}{2}$.

FEUE WHEP
10219

Kinematics

Lorentz invariant phase space ($P \to p_1 + \cdots p_n$)

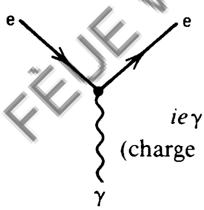
$$dQ = (2\pi)^4 \delta^4 (P - p_1 - \cdots p_n) \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i}$$

Scattering: $$\left.\frac{d\sigma}{d\Omega}\right|_{\text{cm}} = \frac{1}{64\pi^2 s} \frac{p_f}{p_i} |\mathcal{M}|^2$$

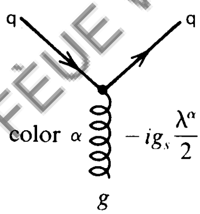
Decay: $$d\Gamma(A \to 1 + \cdots n) = \frac{|\mathcal{M}|^2}{2m_A} dQ.$$

Feynman Rules for $-i\mathcal{M}$

$$\frac{i}{(\not p - m)} \qquad \frac{-ig_{\mu\nu}}{p^2} \qquad \frac{-i\left(g_{\mu\nu} - p_\mu p_\nu / M^2\right)}{p^2 - M^2}$$

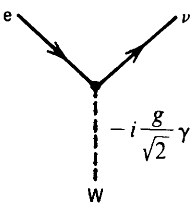


$ie\gamma^\mu$ (charge $-e$)

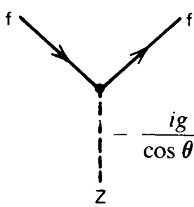


$-ig_s \frac{\lambda^\alpha}{2} \gamma^\mu$

$$\alpha_s = \frac{g_s^2}{4\pi} = \frac{12\pi}{(33 - 2n_f)\log(Q^2/\Lambda^2)}$$



$-i\frac{g}{\sqrt{2}} \gamma^\mu \frac{1}{2}(1 - \gamma^5)$



$-\frac{ig}{\cos\theta_W} \gamma^\mu \frac{1}{2}(c_V^f - c_A^f \gamma^5)$

$$\boxed{\begin{cases} c_V^f = T_f^3 - 2\sin^2\theta_W Q_f \\ c_A^f = T_f^3 \end{cases}}$$

f	Q_f	$(T_f^3)_L$	$(T_f^3)_R$
u, c, t	$+\frac{2}{3}$	$\frac{1}{2}$	0
d, s, b	$-\frac{1}{3}$	$-\frac{1}{2}$	0
ν_e, ν_μ, ν_τ	0	$\frac{1}{2}$	—
e^-, μ^-, τ^-	-1	$-\frac{1}{2}$	0

$$\sin^2\theta_W \simeq 0.23, \; g\sin\theta_W = e, \; G = \frac{\sqrt{2}\, g^2}{8M_W^2} \simeq 1.17 \times 10^{-5}\ \text{GeV}^{-2}$$

FEUE WHEP
10220

Kinematics

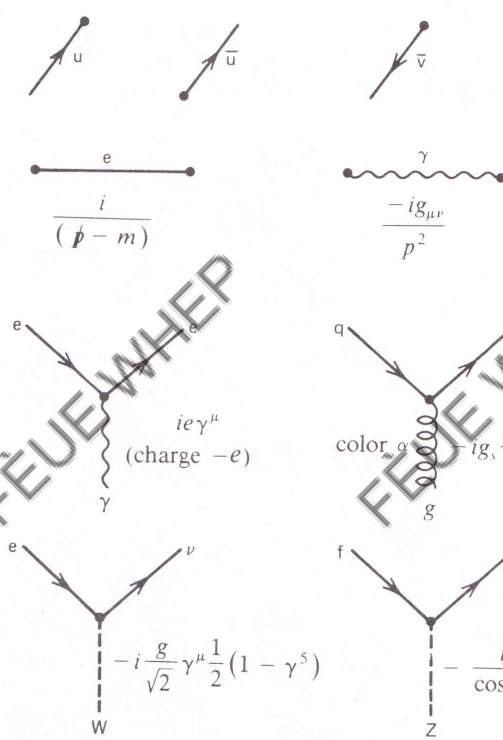
Lorentz invariant phase space ($P \to p_1 + \cdots p$

$$dQ = (2\pi)^4 \delta^4 (P - p_1 - \cdots$$

Scattering: $\left.\frac{d\sigma}{d\Omega}\right|_{\mathrm{cm}} = \frac{1}{64\pi^2 s}\frac{p_f}{p_i}|\mathcal{M}|^2$

Decay: $d\Gamma(A \to 1 + \cdots n) = \frac{|\mathcal{M}|^2}{2m_A}dQ.$

Feynman Rules f



$\frac{i}{(\not{p} - m)}$

$\frac{-ig_{\mu\nu}}{p^2}$

$ie\gamma^\mu$ (charge $-e$)

color α ... $-ig_s$

$-i\frac{g}{\sqrt{2}}\gamma^\mu \frac{1}{2}(1-\gamma^5)$

$-\frac{i}{\cos}$

f
u, c, t
d, s, b
ν_e, ν_μ, ν_τ
e^-, μ^-, τ^-

$\sin^2\theta_W \simeq 0.23,\ g\sin\theta_W$

(continued from front flap)

Assuming a basic knowledge of nonrelativistic quantum mechanics and the theory of special relativity, *Quarks and Leptons* further distinguishes itself with its careful introduction of QED and Feynman rules, and its generalized techniques, which are then applied to QCD and the dynamics of weak interactions. In addition, the text also offers students 200 carefully selected exercises, with their outline solutions collected in the final section of the book.

All in all, *Quarks and Leptons* establishes itself as something of a rarity in the field: it provides the solid 'nuts-and-bolts' approach that an introductory text must have, while it continues to make physics exciting.

About the authors

FRANCIS HALZEN is Professor of Physics at the University of Wisconsin (Madison) where he has taught since 1972. He received his degrees from the University of Leuven in Belgium.

ALAN D. MARTIN is Professor of Theoretical Physics at the University of Durham in England where he has taught since 1964. Dr. Martin is the co-author of *Elementary Particle Theory* with T.D. Spearman. He received his degrees from University College in London.

Both authors have worked at CERN (European Centre for Nuclear Research) in Geneva. Active in elementary particle physics research, they have both been widely published in leading professional journals.

FEUE WHEP
10221

Other titles of interest...

INTRODUCTION TO PLASMA THEORY
Dwight R. Nicholson

Here's the most complete, clearly written introduction to plasma physics available—perfect for senior or first year graduate courses. Students will find *Introduction To Plasma Theory* easier to follow and understand than any other text at this level. No prior knowledge of this topic is assumed—everything is explained carefully and thoroughly. *And none of the mathematical steps are omitted in the derivations.* All of the important topics in plasma theory are given a complete introduction. Instructors may choose to use these introductions as the basis for more in-depth studies of specialized topics. The author also discusses many modern topics—including solitons, parametric instabilities, and weak turbulence theory.
(0 471 09045-X) 1983

INTRODUCTION TO STATISTICAL MECHANICS AND THERMODYNAMICS
Keith Stowe

Intended for a student's very first course in thermodynamics, this introductory textbook uses a statistical approach to the subject while most other texts continue to use the more formal and abstract postulatory method. This statistical approach gives students a broader spectrum of skills, as well as a better understanding of their physical bases. The text includes some 460 exercises and problems, approximately 100 shaded, easily identified summaries, 254 figures, plus a table of physical constants, and a table of symbols used in the text.
(0 471 87058-7) 1983

THEORETICAL MECHANICS
E. Neal Moore

This applications-oriented text presents a thorough, yet highly lucid treatment of classical mechanics at a level suitable for the beginning graduate student. Avoiding the mathematical formalism or specialized approaches of many existing texts, emphasis here is squarely on the solution of physical problems. Throughout, care is taken to *explicitly* show *how* mathematical tools are applied, with a wide variety of illustrative examples included immediately after most theorems.
(0 471 87488-4) 1983

A FIRST COURSE IN ANALYTICAL MECHANICS
Klaus Rossberg

An introduction to classical mechanics emphasizing basic mechanics concepts and their broad application to other areas of physics specifically designed for one semester junior level courses in introductory or analytical mechanics. All topics are systematically developed from the general to the specific, in order to give students insight into the general structure of any physical theory—particularly mechanics. Includes novel problems, lists of procedures, several flow charts, and diagrams showing the relationship between various formalisms of mechanics, and between various physical quantities.
(0 471 86174-X) 1983

JOHN WILEY & SONS, Inc.
605 Third Avenue, New York, N.Y. 10158
New York • Chichester • Brisbane • Toronto • Singa

X001DFG2TP
10222
Used - Good: Quarks and Leptons: An Introductory Course In Modern Particle Physics

Fundamental Mathematics with Proofs

---

1. Arithmetic

Commutative Law of Addition: a + b = b + a - Proof: By Peano axioms or induction.

Associative Law of Addition: (a + b) + c = a + (b + c) - Proof: By induction on natural numbers.

Distributive Law: a(b + c) = ab + a*c - Proof: Multiplication defined as repeated addition.

Exponent Rules: - a^m * a^n = a^(m+n) (Proof: definition of exponents) - (a^m)^n = a^(m*n) (Proof: repeated multiplication)

---

1. Algebra

Identity Elements: a + 0 = a, a * 1 = a - Proof: By definition of 0 and 1.

Inverse Elements: a + (-a) = 0, a * (1/a) = 1 (a != 0) - Proof: By definition of additive/multiplicative inverses.

Solving Equations: - Rules: Add/subtract or multiply/divide both sides by same nonzero number.

Quadratic Formula: x = (-b ± sqrt(b^2 - 4ac))/(2a) - Proof: Completing the square.

---

1. Geometry

Triangle Angle Sum: 180° - Proof: Draw parallel line through vertex, use alternate interior angles.

Pythagorean Theorem: a^2 + b^2 = c^2 - Proof: Using similar triangles or geometric rearrangement.

Area Formulas: - Triangle: 1/2 * base * height - Rectangle: length * width - Circle: pi * r^2

---

1. Trigonometry

Pythagorean Identity: sin^2(x) + cos^2(x) = 1 - Proof: Right triangle definition and unit circle.

Sum/Difference Formulas: - sin(a ± b) = sin(a)cos(b) ± cos(a)sin(b) - cos(a ± b) = cos(a)cos(b) ∓ sin(a)sin(b) - Proof: Using unit circle coordinates.

Law of Sines and Cosines: - Proof: Derived from triangle properties and cosine rule.

---

1. Calculus

Derivative Rules: - Power, sum, product, quotient, chain rules - Proof: Using limit definition of derivative.

Integral Rules: - Sum rule, power rule, substitution rule - Proof: Reverse of derivative, fundamental theorem of calculus.

---

1. Probability & Statistics

Probability Formula: P(E) = favorable/total - Proof: By classical definition.

Sum and Product Rules: - Proof: By counting principle for independent/dependent events.

Mean/Variance: - Proof: Using definition of expectation and squared deviation.

---

1. Number Theory

Even/Odd Rules: - Even + Even = Even, Odd + Odd = Even, Even + Odd = Odd - Proof: By definition: even = 2k, odd = 2k+1

Prime Numbers: - Definition: exactly 2 divisors - Divisibility rules: proven by modular arithmetic

---

1. Set Theory & Logic

Union/Intersection/Complement: - Proofs using element-wise definitions

Logical Operators & Truth Tables: - Proofs by exhaustive enumeration

De Morgan's Laws: - Proof using complements and definitions

10224

MATHEMATICAL METHODS IN THE PHYSICAL SCIENCES

Third Edition

MARY L. BOAS

DePaul University

FÊUE WHEP
10225

MATHEMATICAL METHODS IN THE PHYSICAL SCIENCES

FÊUE WHEP

FÈUE WHEP

MATHEMATICAL METHODS IN THE PHYSICAL SCIENCES

Third Edition

MARY L. BOAS

DePaul University

FEUE WHEP

10228

PUBLISHER	Kaye Pace
SENIOR ACQUISITIONS Editor	Stuart Johnson
PRODUCTION MANAGER	Pam Kennedy
PRODUCTION EDITOR	Sarah Wolfman-Robichaud
MARKETING MANAGER	Amanda Wygal
SENIOR DESIGNER	Dawn Stanley
EDITORIAL ASSISTANT	Krista Jarmas/Alyson Rentrop
PRODUCTION MANAGER	Jan Fisher/Publication Services

This book was set in 10/12 Computer Modern by Publication Services and printed and bound by R.R. Donnelley-Willard. The cover was printed by Lehigh Press.

This book is printed on acid free paper.

Copyright © 2006 John Wiley & Sons, Inc. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978)750-8400, fax (978)750-4470 or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, (201)748-6011, fax (201)748-6008, or online at http://www.wiley.com/go/permissions.

To order books or for customer service please, call 1-800-CALL WILEY (225-5945).

ISBN 0-471-19826-9
ISBN-13 978-0-471-19826-0
ISBN-WIE 0-471-36580-7
ISBN-WIE-13 978-0-471-36580-8

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

10229

To the memory of RPB

10230

FẾUE WHEP

PREFACE

This book is particularly intended for the student with a year (or a year and a half) of calculus who wants to develop, in a short time, a basic competence in each of the many areas of mathematics needed in junior to senior-graduate courses in physics, chemistry, and engineering. Thus it is intended to be accessible to sophomores (or freshmen with AP calculus from high school). It may also be used effectively by a more advanced student to review half-forgotten topics or learn new ones, either by independent study or in a class. Although the book was written especially for students of the physical sciences, students in any field (say mathematics or mathematics for teaching) may find it useful to survey many topics or to obtain some knowledge of areas they do not have time to study in depth. Since theorems are stated carefully, such students should not need to unlearn anything in their later work.

The question of proper mathematical training for students in the physical sciences is of concern to both mathematicians and those who use mathematics in applications. Some instructors may feel that if students are going to study mathematics at all, they should study it in careful and thorough detail. For the undergraduate physics, chemistry, or engineering student, this means either (1) learning more mathematics than a mathematics major or (2) learning a few areas of mathematics thoroughly and the others only from snatches in science courses. The second alternative is often advocated; let me say why I think it is unsatisfactory. It is certainly true that motivation is increased by the immediate application of a mathematical technique, but there are a number of disadvantages:

1. The discussion of the mathematics is apt to be sketchy since that is not the primary concern.
2. Students are faced simultaneously with learning a new mathematical method and applying it to an area of science that is also new to them. Frequently the

FÊUE WHEP
10232

difficulty in comprehending the new scientific area lies more in the distraction caused by poorly understood mathematics than it does in the new scientific ideas.

3. Students may meet what is actually the same mathematical principle in two different science courses without recognizing the connection, or even learn apparently contradictory theorems in the two courses! For example, in thermodynamics students learn that the integral of an exact differential around a closed path is always zero. In electricity or hydrodynamics, they run into $\int_0^{2\pi} d\theta$, which is certainly the integral of an exact differential around a closed path but is not equal to zero!

Now it would be fine if every science student could take the separate mathematics courses in differential equations (ordinary and partial), advanced calculus, linear algebra, vector and tensor analysis, complex variables, Fourier series, probability, calculus of variations, special functions, and so on. However, most science students have neither the time nor the inclination to study that much mathematics, yet they are constantly hampered in their science courses for lack of the basic techniques of these subjects. It is the intent of this book to give these students enough background in each of the needed areas so that they can cope successfully with junior, senior, and beginning graduate courses in the physical sciences. I hope, also, that some students will be sufficiently intrigued by one or more of the fields of mathematics to pursue it futher.

It is clear that something must be omitted if so many topics are to be compressed into one course. I believe that two things can be left out without serious harm at this stage of a student's work: generality, and detailed proofs. Stating and proving a theorem in its most general form is important to the mathematician and to the advanced student, but it is often unnecessary and may be confusing to the more elementary student. This is not in the least to say that science students have no use for careful mathematics. Scientists, even more than pure mathematicians, need careful statements of the limits of applicability of mathematical processes so that they can use them with confidence without having to supply proof of their validity. Consequently I have endeavored to give accurate statements of the needed theorems, although often for special cases or without proof. Interested students can easily find more detail in textbooks in the special fields.

Mathematical physics texts at the senior-graduate level are able to assume a degree of mathematical sophistication and knowledge of advanced physics not yet attained by students at the sophomore level. Yet such students, if given simple and clear explanations, can readily master the techniques we cover in this text. (They not only *can*, but will *have to* in one way or another, if they are going to pass their junior and senior physics courses!) These students are not ready for detailed applications—these they will get in their science courses—but they do need and want to be given some idea of the use of the methods they are studying, and some simple applications. This I have tried to do for each new topic.

For those of you familiar with the second edition, let me outline the changes for the third:

1. Prompted by several requests for matrix diagonalization in Chapter 3, I have moved the first part of Chapter 10 to Chapter 3 and then have amplified the treatment of tensors in Chapter 10. I have also changed Chapter 3 to include more detail about linear vector spaces and then have continued the discussion of basis functions in Chapter 7 (Fourier series), Chapter 8 (Differential equations),

FÊUE WHEP
10233

Chapter 12 (Series solutions) and Chapter 13 (Partial differential equations).

2. Again, prompted by several requests, I have moved Fourier integrals back to the Fourier series Chapter 7. Since this breaks up the integral transforms chapter (old Chapter 15), I decided to abandon that chapter and move the Laplace transform and Dirac delta function material back to the ordinary differential equations Chapter 8. I have also amplified the treatment of the delta function.
3. The Probability chapter (old Chapter 16) now becomes Chapter 15. Here I have changed the title to Probability and Statistics, and have revised the latter part of the chapter to emphasize its purpose, namely to clarify for students the theory behind the rules they learn for handling experimental data.
4. The very rapid development of technological aids to computation poses a steady question for instructors as to their best use. Without selecting any particular Computer Algebra System, I have simply tried for each topic to point out to students both the usefulness and the pitfalls of computer use. (Please see my comments at the end of "To the Student" just ahead.)

The material in the text is so arranged that students who study the chapters in order will have the necessary background at each stage. However, it is not always either necessary or desirable to follow the text order. Let me suggest some rearrangements I have found useful. If students have previously studied the material in any of chapters 1, 3, 4, 5, 6, or 8 (in such courses as second-year calculus, differential equations, linear algebra), then the corresponding chapter(s) could be omitted, used for reference, or, preferably, be reviewed briefly with emphasis on problem solving. Students may know Taylor's theorem, for example, but have little skill in using series approximations; they may know the theory of multiple integrals, but find it difficult to set up a double integral for the moment of inertia of a spherical shell; they may know existence theorems for differential equations, but have little skill in solving, say, $y'' + y = x \sin x$. Problem solving is the essential core of a course on Mathematical Methods.

After Chapters 7 (Fourier Series) and 8 (Ordinary Differential Equations) I like to cover the first four sections of Chapter 13 (Partial Differential Equations). This gives students an introduction to Partial Differential Equations but requires only the use of Fourier series expansions. Later on, after studying Chapter 12, students can return to complete Chapter 13. Chapter 15 (Probability and Statistics) is almost independent of the rest of the text; I have covered this material anywhere from the beginning to the end of a one-year course.

It has been gratifying to hear the enthusiastic responses to the first two editions, and I hope that this third edition will prove even more useful. I want to thank many readers for helpful suggestions and I will appreciate any further comments. If you find misprints, please send them to me at MLBoas@aol.com. I also want to thank the University of Washington physics students who were my LaTeX typists: Toshiko Asai, Jeff Sherman, and Jeffrey Frasca. And I especially want to thank my son, Harold P. Boas, both for mathematical consultations, and for his expert help with LaTeX problems.

Instructors who have adopted the book for a class should consult the publisher about an Instructor's Answer Book, and about a list correlating 2nd and 3rd edition problem numbers for problems which appear in both editions.

Mary L. Boas

FÊUE WHEP

TO THE STUDENT

As you start each topic in this book, you will no doubt wonder and ask "Just why should I study this subject and what use does it have in applications?" There is a story about a young mathematics instructor who asked an older professor "What do you say when students ask about the practical applications of some mathematical topic?" The experienced professor said "I tell them!" This text tries to follow that advice. However, you must on your part be reasonable in your request. It is not possible in one book or course to cover both the mathematical methods and very many detailed applications of them. You will have to be content with some information as to the areas of application of each topic and some of the simpler applications. In your later courses, you will then use these techniques in more advanced applications. At that point you can concentrate on the physical application instead of being distracted by learning new mathematical methods.

One point about your study of this material cannot be emphasized too strongly: To use mathematics effectively in applications, you need not just knowledge but *skill*. Skill can be obtained only through practice. You can obtain a certain superficial *knowledge* of mathematics by listening to lectures, but you cannot obtain *skill* this way. How many students have I heard say "It looks so easy when you do it," or "I understand it but I can't do the problems!" Such statements show lack of practice and consequent lack of skill. The *only* way to develop the skill necessary to use this material in your later courses is to practice by solving many problems. Always study with pencil and paper at hand. Don't just read through a solved problem—try to do it yourself! Then solve some similar ones from the problem set for that section,

FĚUE WHEP
10236

trying to choose the most appropriate method from the solved examples. See the Answers to Selected Problems and check your answers to any problems listed there. If you meet an unfamiliar term, look for it in the Index (or in a dictionary if it is nontechnical).

My students tell me that one of my most frequent comments to them is "You're working too hard." There is no merit in spending hours producing a solution to a problem that can be done by a better method in a few minutes. Please ignore anyone who disparages problem-solving techniques as "tricks" or "shortcuts." You will find that the more able you are to choose effective methods of solving problems in your science courses, the easier it will be for you to master new material. But this means practice, practice, practice! The *only* way to learn to solve problems is to solve problems. In this text, you will find both drill problems and harder, more challenging problems. You should not feel satisfied with your study of a chapter until you can solve a reasonable number of these problems.

You may be thinking "I don't really need to study this—my computer will solve all these problems for me." Now Computer Algebra Systems are wonderful—as you know, they save you a lot of laborious calculation and quickly plot graphs which clarify a problem. But a computer is a tool; *you* are the one in charge. A very perceptive student recently said to me (about the use of a computer for a special project): "*First* you learn how to do it; *then* you see what the computer can do to make it easier." Quite so! A very effective way to study a new technique is to do some simple problems by hand in order to understand the process, and compare your results with a computer solution. You will then be better able to use the method to set up and solve similar more complicated applied problems in your advanced courses. So, in one problem set after another, I will remind you that the point of solving some simple problems is not to get an answer (which a computer will easily supply) but rather to learn the ideas and techniques which will be so useful in your later courses.

M. L. B.

10237

CONTENTS

FÊUE WHEP
10238

FEUE WHEP

FEUE WHEP
10240

10241

FEUE WHEP
10242

10243

CHAPTER 1

Infinite Series, Power Series

1. THE GEOMETRIC SERIES

As a simple example of many of the ideas involved in series, we are going to consider the geometric series. You may recall that in a geometric progression we multiply each term by some fixed number to get the next term. For example, the *sequences*

$$2,\ 4,\ 8,\ 16,\ 32, \ldots, \tag{1.1a}$$

$$1,\ \tfrac{2}{3},\ \tfrac{4}{9},\ \tfrac{8}{27},\ \tfrac{16}{81}, \ldots, \tag{1.1b}$$

$$a,\ ar,\ ar^2,\ ar^3, \ldots, \tag{1.1c}$$

are geometric progressions. It is easy to think of examples of such progressions. Suppose the number of bacteria in a culture doubles every hour. Then the terms of (1.1a) represent the number by which the bacteria population has been multiplied after 1 hr, 2 hr, and so on. Or suppose a bouncing ball rises each time to $\frac{2}{3}$ of the height of the previous bounce. Then (1.1b) would represent the heights of the successive bounces in yards if the ball is originally dropped from a height of 1 yd.

In our first example it is clear that the bacteria population would increase without limit as time went on (mathematically, anyway; that is, assuming that nothing like lack of food prevented the assumed doubling each hour). In the second example, however, the height of bounce of the ball decreases with successive bounces, and we might ask for the total distance the ball goes. The ball falls a distance 1 yd, rises a distance $\frac{2}{3}$ yd and falls a distance $\frac{2}{3}$ yd, rises a distance $\frac{4}{9}$ yd and falls a distance $\frac{4}{9}$ yd, and so on. Thus it seems reasonable to write the following expression for the total distance the ball goes:

$$1 + 2\cdot\tfrac{2}{3} + 2\cdot\tfrac{4}{9} + 2\cdot\tfrac{8}{27} + \cdots = 1 + 2\left(\tfrac{2}{3} + \tfrac{4}{9} + \tfrac{8}{27} + \cdots\right), \tag{1.2}$$

where the three dots mean that the terms continue as they have started (each one being $\frac{2}{3}$ the preceding one), and there is never a last term. Let us consider the expression in parentheses in (1.2), namely

$$\frac{2}{3} + \frac{4}{9} + \frac{8}{27} + \cdots. \tag{1.3}$$

FEUE WHEP
10244

This expression is an example of an *infinite series*, and we are asked to find its sum. Not all infinite series have sums; you can see that the series formed by adding the terms in (1.1a) does not have a finite sum. However, even when an infinite series does have a finite sum, we cannot find it by adding the terms because no matter how many we add there are always more. Thus we must find another method. (It is actually deeper than this; what we really have to do is to *define* what we mean by the sum of the series.)

Let us first find the sum of n terms in (1.3). The formula (Problem 2) for the sum of n terms of the geometric progression (1.1c) is

$$S_n = \frac{a(1-r^n)}{1-r}. \tag{1.4}$$

Using (1.4) in (1.3), we find

$$S_n = \frac{2}{3} + \frac{4}{9} + \cdots + \left(\frac{2}{3}\right)^n = \frac{\frac{2}{3}[1-(\frac{2}{3})^n]}{1-\frac{2}{3}} = 2\left[1-\left(\frac{2}{3}\right)^n\right]. \tag{1.5}$$

As n increases, $(\frac{2}{3})^n$ decreases and approaches zero. Then the sum of n terms approaches 2 as n increases, and we say that the sum of the series is 2. (This is really a definition: The sum of an infinite series is the limit of the sum of n terms as $n \to \infty$.) Then from (1.2), the total distance traveled by the ball is $1 + 2 \cdot 2 = 5$. This is an answer to a mathematical problem. A physicist might well object that a bounce the size of an atom is nonsense! However, after a number of bounces, the remaining infinite number of small terms contribute very little to the final answer (see Problem 1). Thus it makes little difference (in our answer for the total distance) whether we insist that the ball rolls after a certain number of bounces or whether we include the entire series, and it is easier to find the sum of the series than to find the sum of, say, twenty terms.

Series such as (1.3) whose terms form a geometric progression are called *geometric series*. We can write a geometric series in the form

$$a + ar + ar^2 + \cdots + ar^{n-1} + \cdots. \tag{1.6}$$

The sum of the geometric series (if it has one) is by definition

$$S = \lim_{n\to\infty} S_n, \tag{1.7}$$

where S_n is the sum of n terms of the series. By following the method of the example above, you can show (Problem 2) that a geometric series has a sum if and only if $|r| < 1$, and in this case the sum is

$$S = \frac{a}{1-r}. \tag{1.8}$$

FEUE WHEP

The series is then called *convergent*.

Here is an interesting use of (1.8). We can write $0.3333\cdots = \frac{3}{10} + \frac{3}{100} + \frac{3}{1000} + \cdots = \frac{3/10}{1-1/10} = \frac{1}{3}$ by (1.8). Now of course you knew that, but how about $0.785714285714\cdots$? We can write this as $0.5+0.285714285714\cdots = \frac{1}{2} + \frac{0.285714}{1-10^{-6}} = \frac{1}{2} + \frac{285714}{999999} = \frac{1}{2} + \frac{2}{7} = \frac{11}{14}$. (Note that any repeating decimal is equivalent to a fraction which can be found by this method.) If you want to use a computer to do the arithmetic, be sure to tell it to give you an exact answer or it may hand you back the decimal you started with! You can also use a computer to sum the series, but using (1.8) may be simpler. (Also see Problem 14.)

PROBLEMS, SECTION 1

1. In the bouncing ball example above, find the height of the tenth rebound, and the distance traveled by the ball after it touches the ground the tenth time. Compare this distance with the total distance traveled.

2. Derive the formula (1.4) for the sum S_n of the geometric progression $S_n = a + ar + ar^2 + \cdots + ar^{n-1}$. *Hint:* Multiply S_n by r and subtract the result from S_n; then solve for S_n. Show that the geometric series (1.6) converges if and only if $|r| < 1$; also show that if $|r| < 1$, the sum is given by equation (1.8).

Use equation (1.8) to find the fractions that are equivalent to the following repeating decimals:

3. $0.55555\cdots$ **4.** $0.818181\cdots$ **5.** $0.583333\cdots$

6. $0.61111\cdots$ **7.** $0.185185\cdots$ **8.** $0.694444\cdots$

9. $0.857142857142\cdots$ **10.** $0.576923076923076923\cdots$

11. $0.678571428571428571\cdots$

12. In a water purification process, one-nth of the impurity is removed in the first stage. In each succeeding stage, the amount of impurity removed is one-nth of that removed in the preceding stage. Show that if $n = 2$, the water can be made as pure as you like, but that if $n = 3$, at least one-half of the impurity will remain no matter how many stages are used.

13. If you invest a dollar at "6% interest compounded monthly," it amounts to $(1.005)^n$ dollars after n months. If you invest \$10 at the beginning of each month for 10 years (120 months), how much will you have at the end of the 10 years?

14. A computer program gives the result 1/6 for the sum of the series $\sum_{n=0}^{\infty}(-5)^n$. Show that this series is divergent. Do you see what happened? *Warning hint:* Always consider whether an answer is reasonable, whether it's a computer answer or your work by hand.

15. Connect the midpoints of the sides of an equilateral triangle to form 4 smaller equilateral triangles. Leave the middle small triangle blank, but for each of the other 3 small triangles, draw lines connecting the midpoints of the sides to create 4 tiny triangles. Again leave each middle tiny triangle blank and draw the lines to divide the others into 4 parts. Find the infinite series for the total area left blank if this process is continued indefinitely. (Suggestion: Let the area of the original triangle be 1; then the area of the first blank triangle is 1/4.) Sum the series to find the total area left blank. Is the answer what you expect? *Hint:* What is the "area" of a straight line? (Comment: You have constructed a *fractal* called the Sierpiński gasket. A fractal has the property that a magnified view of a small part of it looks very much like the original.)

FEUE WHEP

16. Suppose a large number of particles are bouncing back and forth between $x = 0$ and $x = 1$, except that at each endpoint some escape. Let r be the fraction reflected each time; then $(1 - r)$ is the fraction escaping. Suppose the particles start at $x = 0$ heading toward $x = 1$; eventually all particles will escape. Write an infinite series for the fraction which escape at $x = 1$ and similarly for the fraction which escape at $x = 0$. Sum both the series. What is the largest fraction of the particles which can escape at $x = 0$? (Remember that r must be between 0 and 1.)

2. DEFINITIONS AND NOTATION

There are many other infinite series besides geometric series. Here are some examples:

$$1^2 + 2^2 + 3^2 + 4^2 + \cdots, \tag{2.1a}$$

$$\frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \frac{4}{2^4} + \cdots, \tag{2.1b}$$

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots. \tag{2.1c}$$

In general, an infinite series means an expression of the form

$$a_1 + a_2 + a_3 + \cdots + a_n + \cdots, \tag{2.2}$$

where the a_n's (one for each positive integer n) are numbers or functions given by some formula or rule. The three dots in each case mean that the series never ends. The terms continue according to the law of formation, which is supposed to be evident to you by the time you reach the three dots. If there is apt to be doubt about how the terms are formed, a general or nth term is written like this:

$$1^2 + 2^2 + 3^2 + \cdots + n^2 + \cdots, \tag{2.3a}$$

$$x - x^2 + \frac{x^3}{2} + \cdots + \frac{(-1)^{n-1}x^n}{(n-1)!} + \cdots. \tag{2.3b}$$

(The quantity $n!$, read n factorial, means, for integral n, the product of all integers from 1 to n; for example, $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$. The quantity $0!$ is defined to be 1.) In (2.3a), it is easy to see without the general term that each term is just the square of the number of the term, that is, n^2. However, in (2.3b), if the formula for the general term were missing, you could probably make several reasonable guesses for the next term. To be sure of the law of formation, we must either know a good many more terms or have the formula for the general term. You should verify that the fourth term in (2.3b) is $-x^4/6$.

We can also write series in a shorter abbreviated form using a summation sign $\sum$ followed by the formula for the nth term. For example, (2.3a) would be written

$$1^2 + 2^2 + 3^2 + 4^2 + \cdots = \sum_{n=1}^{\infty} n^2 \tag{2.4}$$

(read "the sum of n^2 from $n = 1$ to ∞"). The series (2.3b) would be written

$$x - x^2 + \frac{x^3}{2} - \frac{x^4}{6} + \cdots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}x^n}{(n-1)!}$$

10247

For printing convenience, sums like (2.4) are often written $\sum_{n=1}^{\infty} n^2$.

In Section 1, we have mentioned both sequences and series. The lists in (1.1) are sequences; a *sequence* is simply a set of quantities, one for each n. A *series* is an indicated sum of such quantities, as in (1.3) or (1.6). We will be interested in various sequences related to a series: for example, the sequence a_n of terms of the series, the sequence S_n of partial sums [see (1.5) and (4.5)], the sequence R_n [see (4.7)], and the sequence ρ_n [see (6.2)]. In all these examples, we want to find the limit of a sequence as $n \to \infty$ (if the sequence has a limit). Although limits can be found by computer, many simple limits can be done faster by hand.

Example 1. Find the limit as $n \to \infty$ of the sequence

$$\frac{(2n-1)^4 + \sqrt{1+9n^8}}{1-n^3-7n^4}.$$

We divide numerator and denominator by n^4 and take the limit as $n \to \infty$. Then all terms go to zero except

$$\frac{2^4+\sqrt{9}}{-7} = -\frac{19}{7}.$$

Example 2. Find $\lim_{n\to\infty} \frac{\ln n}{n}$. By L'Hôpital's rule (see Section 15)

$$\lim_{n\to\infty} \frac{\ln n}{n} = \lim_{n\to\infty} \frac{1/n}{1} = 0.$$

Comment: Strictly speaking, we can't differentiate a function of n if n is an integer, but we can consider $f(x) = (\ln x)/x$, and the limit of the sequence is the same as the limit of $f(x)$.

Example 3. Find $\lim_{n\to\infty} \left(\frac{1}{n}\right)^{1/n}$. We first find

$$\ln\left(\frac{1}{n}\right)^{1/n} = -\frac{1}{n}\ln n.$$

Then by Example 2, the limit of $(\ln n)/n$ is 0, so the original limit is $e^0 = 1$.

PROBLEMS, SECTION 2

In the following problems, find the limit of the given sequence as $n \to \infty$.

1. $\dfrac{n^2+5n^3}{2n^3+3\sqrt{4+n^6}}$ **2.** $\dfrac{(n+1)^2}{\sqrt{3+5n^2+4n^4}}$ **3.** $\dfrac{(-1)^n\sqrt{n+1}}{n}$

4. $\dfrac{2^n}{n^2}$ **5.** $\dfrac{10^n}{n!}$ **6.** $\dfrac{n^n}{n!}$

7. $(1+n^2)^{1/\ln n}$ **8.** $\dfrac{(n!)^2}{(2n)!}$ **9.** $n\sin(1/n)$

10248

3. APPLICATIONS OF SERIES

In the example of the bouncing ball in Section 1, we saw that it is possible for the sum of an infinite series to be nearly the same as the sum of a fairly small number of terms at the beginning of the series (also see Problem 1.1). Many applied problems cannot be solved exactly, but we may be able to find an answer in terms of an infinite series, and then use only as many terms as necessary to obtain the needed accuracy. We shall see many examples of this both in this chapter and in later chapters. Differential equations (see Chapters 8 and 12) and partial differential equations (see Chapter 13) are frequently solved by using series. We will learn how to find series that represent functions; often a complicated function can be approximated by a few terms of its series (see Section 15).

But there is more to the subject of infinite series than making approximations. We will see (Chapter 2, Section 8) how we can use power series (that is, series whose terms are powers of x) to give meaning to functions of complex numbers, and (Chapter 3, Section 6) how to define a function of a matrix using the power series of the function. Also power series are just a first example of infinite series. In Chapter 7 we will learn about Fourier series (whose terms are sines and cosines). In Chapter 12, we will use power series to solve differential equations, and in Chapters 12 and 13, we will discuss other series such as Legendre and Bessel. Finally, in Chapter 14, we will discover how a study of power series clarifies our understanding of the mathematical functions we use in applications.

4. CONVERGENT AND DIVERGENT SERIES

We have been talking about series which have a finite sum. We have also seen that there are series which do not have finite sums, for example (2.1a). If a series has a finite sum, it is called *convergent*. Otherwise it is called *divergent*. It is important to know whether a series is convergent or divergent. Some weird things can happen if you try to apply ordinary algebra to a divergent series. Suppose we try it with the following series:

$$S = 1 + 2 + 4 + 8 + 16 + \cdots . \tag{4.1}$$

Then,

$$\begin{aligned} 2S &= 2 + 4 + 8 + 16 + \cdots = S - 1, \\ S &= -1. \end{aligned}$$

This is obvious nonsense, and you may laugh at the idea of trying to operate with such a violently divergent series as (4.1). But the same sort of thing can happen in more concealed fashion, and has happened and given wrong answers to people who were not careful enough about the way they used infinite series. At this point you probably would not recognize that the series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots \tag{4.2}$$

is divergent, but it is; and the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots \tag{4.3}$$

FÊUE WHEP
10249

is convergent as it stands, but can be made to have *any* sum you like by combining the terms in a different order! (See Section 8.) You can see from these examples how essential it is to know whether a series converges, and also to know how to apply algebra to series correctly. There are even cases in which some divergent series can be used (see Chapter 11), but in this chapter we shall be concerned with convergent series.

Before we consider some tests for convergence, let us repeat the definition of convergence more carefully. Let us call the terms of the series a_n so that the series is

$$a_1 + a_2 + a_3 + a_4 + \cdots + a_n + \cdots . \tag{4.4}$$

Remember that the three dots mean that there is never a last term; the series goes on without end. Now consider the sums S_n that we obtain by adding more and more terms of the series. We define

$$\begin{aligned} S_1 &= a_1, \\ S_2 &= a_1 + a_2, \\ S_3 &= a_1 + a_2 + a_3, \\ &\cdots \\ S_n &= a_1 + a_2 + a_3 + \cdots + a_n. \end{aligned} \tag{4.5}$$

Each S_n is called a *partial sum*; it is the sum of the first n terms of the series. We had an example of this for a geometric progression in (1.4). The letter n can be any integer; for each n, S_n stops with the nth term. (Since S_n is not an infinite series, there is no question of convergence for it.) As n increases, the partial sums may increase without any limit as in the series (2.1a). They may oscillate as in the series $1-2+3-4+5-\cdots$ (which has partial sums $1, -1, 2, -2, 3, \cdots$) or they may have some more complicated behavior. One possibility is that the S_n's may, after a while, not change very much any more; the a_n's may become very small, and the S_n's come closer and closer to some value S. We are particularly interested in this case in which the S_n's approach a limiting value, say

$$\lim_{n\to\infty} S_n = S. \tag{4.6}$$

(It is understood that S is a finite number.) If this happens, we make the following definitions.

a. If the partial sums S_n of an infinite series tend to a limit S, the series is called *convergent*. Otherwise it is called *divergent*.

b. The limiting value S is called the *sum of the series*.

c. The difference $R_n = S - S_n$ is called the *remainder* (or the remainder after n terms). From (4.6), we see that

$$\lim_{n\to\infty} R_n = \lim_{n\to\infty} (S - S_n) = S - S = 0. \tag{4.7}$$

10250

▶ **Example 1.** We have already (Section 1) found S_n and S for a geometric series. From (1.8) and (1.4), we have for a geometric series, $R_n = \frac{ar^n}{1-r}$ which $\to 0$ as $n \to \infty$ if $|r| < 1$.

▶ **Example 2.** By partial fractions, we can write $\frac{2}{n^2-1} = \frac{1}{n-1} - \frac{1}{n+1}$. Let's write out a number of terms of the series

$$\sum_2^\infty \frac{2}{n^2-1} = \sum_2^\infty \left(\frac{1}{n-1} - \frac{1}{n+1}\right) = \sum_1^\infty \left(\frac{1}{n} - \frac{1}{n+2}\right)$$
$$= 1 - \frac{1}{3} + \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{5} + \frac{1}{4} - \frac{1}{6} + \frac{1}{5} - \frac{1}{7} + \frac{1}{6} - \frac{1}{8} + \cdots$$
$$+ \frac{1}{n-2} - \frac{1}{n} + \frac{1}{n-1} - \frac{1}{n+1} + \frac{1}{n} - \frac{1}{n+2} + \cdots.$$

Note the cancellation of terms; this kind of series is called a telescoping series. Satisfy yourself that when we have added the nth term $(\frac{1}{n} - \frac{1}{n+2})$, the only terms which have not cancelled are $1, \frac{1}{2}, \frac{-1}{n+1}$, and $\frac{-1}{n+2}$, so we have

$$S_n = \frac{3}{2} - \frac{1}{n+1} - \frac{1}{n+2}, \qquad S = \frac{3}{2}, \qquad R_n = \frac{1}{n+1} + \frac{1}{n+2}.$$

▶ **Example 3.** Another interesting series is

$$\sum_1^\infty \ln\left(\frac{n}{n+1}\right) = \sum_1^\infty [\ln n - \ln(n+1)]$$
$$= \ln 1 - \ln 2 + \ln 2 - \ln 3 + \ln 3 - \ln 4 + \cdots + \ln n - \ln(n+1) \cdots.$$

Then $S_n = -\ln(n+1)$ which $\to -\infty$ as $n \to \infty$, so the series diverges. However, note that $a_n = \ln \frac{n}{n+1} \to \ln 1 = 0$ as $n \to \infty$, so we see that even if the terms tend to zero, a series may diverge.

▶ PROBLEMS, SECTION 4

For the following series, write formulas for the sequences a_n, S_n, and R_n, and find the limits of the sequences as $n \to \infty$ (if the limits exist).

1. $\displaystyle\sum_1^\infty \frac{1}{2^n}$

2. $\displaystyle\sum_0^\infty \frac{1}{5^n}$

3. $1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} \cdots$

4. $\displaystyle\sum_1^\infty e^{-n \ln 3}$ *Hint:* What is $e^{-\ln 3}$?

5. $\displaystyle\sum_0^\infty e^{2n \ln \sin(\pi/3)}$ *Hint:* Simplify this.

6. $\displaystyle\sum_1^\infty \frac{1}{n(n+1)}$ *Hint:* $\dfrac{1}{n(n+1)} = \dfrac{1}{n} - \dfrac{1}{n+1}$.

7. $\dfrac{3}{1 \cdot 2} - \dfrac{5}{2 \cdot 3} + \dfrac{7}{3 \cdot 4} - \dfrac{9}{4 \cdot 5} + \cdots$

FÊUE WHEP

5. TESTING SERIES FOR CONVERGENCE; THE PRELIMINARY TEST

It is not in general possible to write a simple formula for S_n and find its limit as $n \to \infty$ (as we have done for a few special series), so we need some other way to find out whether a given series converges. Here we shall consider a few simple tests for convergence. These tests will illustrate some of the ideas involved in testing series for convergence and will work for a good many, but not all, cases. There are more complicated tests which you can find in other books. In some cases it may be quite a difficult mathematical problem to investigate the convergence of a complicated series. However, for our purposes the simple tests we give here will be sufficient.

First we discuss a useful *preliminary test.* In most cases you should apply this to a series before you use other tests.

> **Preliminary test.** If the terms of an infinite series do *not* tend to zero (that is, if $\lim_{n\to\infty} a_n \neq 0$), the series diverges. If $\lim_{n\to\infty} a_n = 0$, we must test further.

This is *not* a test for convergence; what it does is to weed out some very badly divergent series which you then do not have to spend time testing by more complicated methods. *Note carefully:* The preliminary test can *never* tell you that a series converges. It does *not* say that series converge if $a_n \to 0$ and, in fact, often they do not. A simple example is the harmonic series (4.2); the nth term certainly tends to zero, but we shall soon show that the series $\sum_{n=1}^{\infty} 1/n$ is divergent. On the other hand, in the series

$$\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \frac{4}{5} + \cdots$$

the terms are tending to 1, so by the preliminary test, this series diverges and no further testing is needed.

PROBLEMS, SECTION 5

Use the preliminary test to decide whether the following series are divergent or require further testing. *Careful:* Do *not* say that a series is convergent; the preliminary test cannot decide this.

1. $\frac{1}{2} - \frac{4}{5} + \frac{9}{10} - \frac{16}{17} + \frac{25}{26} - \frac{36}{37} + \cdots$

2. $\sqrt{2} + \frac{\sqrt{3}}{2} + \frac{\sqrt{4}}{3} + \frac{\sqrt{5}}{4} + \frac{\sqrt{6}}{5} + \cdots$

3. $\sum_{n=1}^{\infty} \frac{n+3}{n^2+10n}$

4. $\sum_{n=1}^{\infty} \frac{(-1)^n n^2}{(n+1)^2}$

5. $\sum_{n=1}^{\infty} \frac{n!}{n!+1}$

6. $\sum_{n=1}^{\infty} \frac{n!}{(n+1)!}$

7. $\sum_{n=1}^{\infty} \frac{(-1)^n n}{\sqrt{n^3+1}}$

8. $\sum_{n=1}^{\infty} \frac{\ln n}{n}$

9. $\sum_{n=1}^{\infty} \frac{3^n}{2^n+3^n}$

10. $\sum_{n=2}^{\infty} \left(1 - \frac{1}{n^2}\right)$

11. Using (4.6), give a proof of the preliminary test. *Hint:* $S_n - S_{n-1} = a_n$.

FEUE WHEP

6. CONVERGENCE TESTS FOR SERIES OF POSITIVE TERMS; ABSOLUTE CONVERGENCE

We are now going to consider four useful tests for series whose terms are all positive. If some of the terms of a series are negative, we may still want to consider the related series which we get by making all the terms positive; that is, we may consider the series whose terms are the absolute values of the terms of our original series. If this new series converges, we call the original series *absolutely convergent*. It can be proved that if a series converges absolutely, then it converges (Problem 7.9). This means that if the series of absolute values converges, the series is still convergent when you put back the original minus signs. (The sum is different, of course.) The following four tests may be used, then, either for testing series of positive terms, or for testing any series for absolute convergence.

A. The Comparison Test

This test has two parts, (a) and (b).

(a) Let

$$m_1 + m_2 + m_3 + m_4 + \cdots$$

be a series of positive terms which you know converges. Then the series you are testing, namely

$$a_1 + a_2 + a_3 + a_4 + \cdots$$

is absolutely convergent if $|a_n| \le m_n$ (that is, if the absolute value of each term of the a series is no larger than the corresponding term of the m series) for all n from some point on, say after the third term (or the millionth term). See the example and discussion below.

(b) Let

$$d_1 + d_2 + d_3 + d_4 + \cdots$$

be a series of positive terms which you know diverges. Then the series

$$|a_1| + |a_2| + |a_3| + |a_4| + \cdots$$

diverges if $|a_n| \ge d_n$ for all n from some point on.

Warning: Note carefully that neither $|a_n| \ge m_n$ nor $|a_n| \le d_n$ tells us anything. That is, if a series has terms larger than those of a convergent series, it may still converge or it may diverge—we must test it further. Similarly, if a series has terms smaller than those of a divergent series, it may still diverge, or it may converge.

▶ **Example.** Test $\sum_{n=1}^{\infty} \frac{1}{n!} = 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \cdots$ for convergence.

As a comparison series, we choose the geometric series

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots .$$

Notice that we do not care about the first few terms (or, in fact, any finite number of terms) in a series, because they can affect the sum of the series but *not* whether

10253

it converges. When we ask whether a series converges or not, we are asking what happens as we add more and more terms for larger and larger n. Does the sum increase indefinitely, or does it approach a limit? What the first five or hundred or million terms are has no effect on whether the sum eventually increases indefinitely or approaches a limit. Consequently we frequently ignore some of the early terms in testing series for convergence.

In our example, the terms of $\sum_{n=1}^{\infty} 1/n!$ are smaller than the corresponding terms of $\sum_{n=1}^{\infty} 1/2^n$ for all $n > 3$ (Problem 1). We know that the geometric series converges because its ratio is $\frac{1}{2}$. Therefore $\sum_{n=1}^{\infty} 1/n!$ converges also.

PROBLEMS, SECTION 6

1. Show that $n! > 2^n$ for all $n > 3$. *Hint:* Write out a few terms; then consider what you multiply by to go from, say, $5!$ to $6!$ and from 2^5 to 2^6.

2. Prove that the harmonic series $\sum_{n=1}^{\infty} 1/n$ is divergent by comparing it with the series

$$1+\frac{1}{2}+\left(\frac{1}{4}+\frac{1}{4}\right)+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)+\left(8 \text{ terms each equal to } \frac{1}{16}\right)+\cdots,$$

which is $\quad 1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\cdots.$

3. Prove the convergence of $\sum_{n=1}^{\infty} 1/n^2$ by grouping terms somewhat as in Problem 2.

4. Use the comparison test to prove the convergence of the following series:

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2^n+3^n}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n\,2^n}$

5. Test the following series for convergence using the comparison test.

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ *Hint:* Which is larger, n or $\sqrt{n}$? (b) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{\ln n}$

6. There are 9 one-digit numbers (1 to 9), 90 two-digit numbers (10 to 99). How many three-digit, four-digit, etc., numbers are there? The first 9 terms of the harmonic series $1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{9}$ are all greater than $\frac{1}{10}$; similarly consider the next 90 terms, and so on. Thus prove the divergence of the harmonic series by comparison with the series

$$\left[\tfrac{1}{10}+\tfrac{1}{10}+\cdots\left(9 \text{ terms each } = \tfrac{1}{10}\right)\right]+\left[\ 90 \text{ terms each } = \tfrac{1}{100}\right]+\cdots$$
$$= \tfrac{9}{10}+\tfrac{90}{100}+\cdots = \tfrac{9}{10}+\tfrac{9}{10}+\cdots.$$

The comparison test is really the basic test from which other tests are derived. It is probably the most useful test of all for the experienced mathematician but it is often hard to think of a satisfactory m series until you have had a good deal of experience with series. Consequently, you will probably not use it as often as the next three tests.

B. The Integral Test

We can use this test when the terms of the series are positive and not increasing, that is, when $a_{n+1} \le a_n$. (Again remember that we can ignore any finite number of terms of the series; thus the test can still be used even if the condition $a_{n+1} \le a_n$ does not hold for a finite number of terms.) To apply the test we think of a_n as a

FÉUE WHEP 10254

function of the variable n, and, forgetting our previous meaning of n, we allow it to take all values, not just integral ones. The test states that:

> If $0 < a_{n+1} \le a_n$ for $n > N$, then $\sum^{\infty} a_n$ converges if $\int^{\infty} a_n\,dn$ is finite and diverges if the integral is infinite. (The integral is to be evaluated *only* at the upper limit; no lower limit is needed.)

To understand this test, imagine a graph sketched of a_n as a function of n. For example, in testing the harmonic series $\sum_{n=1}^{\infty} 1/n$, we consider the graph of the function $y = 1/n$ (similar to Figures 6.1 and 6.2) letting n have all values, not just integral ones. Then the values of y on the graph at $n = 1, 2, 3, \cdots$, are the terms of the series. In Figures 6.1 and 6.2, the areas of the rectangles are just the terms of the series. Notice that in Figure 6.1 the top edge of each rectangle is above the curve, so that the area of the rectangles is greater than the corresponding area under the curve. On the other hand, in Figure 6.2 the rectangles lie below the curve, so their area is less than the corresponding area under the curve. Now the areas of the rectangles are just the terms of the series, and the area under the curve is an integral of $y\,dn$ or $a_n\,dn$. The upper limit on the integrals is ∞ and the lower limit could be made to correspond to any term of the series we wanted to start with. For example (see Figure 6.1), $\int_3^{\infty} a_n\,dn$ is less than the sum of the series from a_3 on, but (see Figure 6.2) greater than the sum of the series from a_4 on. If the integral is finite, then the sum of the series from a_4 on is finite, that is, the series converges. Note again that the terms at the beginning of a series have nothing to do with convergence. On the other hand, if the integral is infinite, then the sum of the series from a_3 on is infinite and the series diverges. Since the beginning terms are of no interest, you should simply evaluate $\int^{\infty} a_n\,dn$. (Also see Problem 16.)

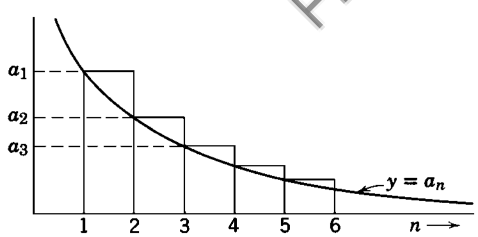


Figure 6.1

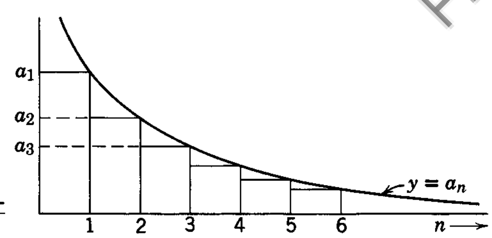


Figure 6.2

▶ **Example.** Test for convergence the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots. \tag{6.1}$$

Using the integral test, we evaluate

$$\int^{\infty} \frac{1}{n}\,dn = \ln n\Big|^{\infty} = \infty.$$

(We use the symbol ln to mean a natural logarithm, that is, a logarithm to the base e.) Since the integral is infinite, the series diverges.

10255

PROBLEMS, SECTION 6

Use the integral test to find whether the following series converge or diverge. *Hint* and *warning:* Do *not* use lower limits on your integrals (see Problem 16).

7. $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n \ln n}$ **8.** $\displaystyle\sum_{n=1}^{\infty} \frac{n}{n^2+4}$ **9.** $\displaystyle\sum_{n=3}^{\infty} \frac{1}{n^2-4}$

10. $\displaystyle\sum_{n=1}^{\infty} \frac{e^n}{e^{2n}+9}$ **11.** $\displaystyle\sum_{1}^{\infty} \frac{1}{n(1+\ln n)^{3/2}}$ **12.** $\displaystyle\sum_{1}^{\infty} \frac{n}{(n^2+1)^2}$

13. $\displaystyle\sum_{1}^{\infty} \frac{n^2}{n^3+1}$ **14.** $\displaystyle\sum_{1}^{\infty} \frac{1}{\sqrt{n^2+9}}$

15. Use the integral test to prove the following so-called p-series test. The series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \quad \text{is} \quad \begin{cases} \text{convergent} & \text{if } p > 1, \\ \text{divergent} & \text{if } p \le 1. \end{cases}$$

Caution: Do $p = 1$ separately.

16. In testing $\sum 1/n^2$ for convergence, a student evaluates $\int_0^\infty n^{-2} dn = -n^{-1}|_0^\infty = 0 + \infty = \infty$ and concludes (erroneously) that the series diverges. What is wrong? *Hint:* Consider the area under the curve in a diagram such as Figure 6.1 or 6.2. This example shows the danger of using a lower limit in the integral test.

17. Use the integral test to show that $\sum_{n=0}^{\infty} e^{-n^2}$ converges. *Hint:* Although you cannot *evaluate* the integral, you can show that it is finite (which is all that is necessary) by comparing it with $\int^\infty e^{-n} dn$.

C. The Ratio Test

The integral test depends on your being able to integrate $a_n dn$; this is not always easy! We consider another test which will handle many cases in which we cannot evaluate the integral. Recall that in the geometric series each term could be obtained by multiplying the one before it by the ratio r, that is, $a_{n+1} = ra_n$ or $a_{n+1}/a_n = r$. For other series the ratio a_{n+1}/a_n is not constant but depends on n; let us call the absolute value of this ratio ρ_n. Let us also find the limit (if there is one) of the sequence ρ_n as $n \to \infty$ and call this limit ρ. Thus we define ρ_n and ρ by the equations

$$\rho_n = \left|\frac{a_{n+1}}{a_n}\right|, \qquad \rho = \lim_{n\to\infty} \rho_n. \tag{6.2}$$

If you recall that a geometric series converges if $|r| < 1$, it may seem plausible that a series with $\rho < 1$ should converge and this is true. This statement can be proved (Problem 30) by comparing the series to be tested with a geometric series. Like a geometric series with $|r| > 1$, a series with $\rho > 1$ also diverges (Problem 30). However, if $\rho = 1$, the ratio test does not tell us anything; some series with $\rho = 1$ converge

FEUE WHEP
10256

and some diverge, so we must find another test (say one of the two preceding tests). To summarize the ratio test:

$$\text{If } \begin{cases} \rho < 1, & \text{the series converges;} \\ \rho = 1, & \text{use a different test;} \\ \rho > 1, & \text{the series diverges.} \end{cases} \tag{6.3}$$

▶ **Example 1.** Test for convergence the series

$$1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots.$$

Using (6.2), we have

$$\begin{aligned} \rho_n &= \left| \frac{1}{(n+1)!} \div \frac{1}{n!} \right| \\ &= \frac{n!}{(n+1)!} = \frac{n(n-1)\cdots 3 \cdot 2 \cdot 1}{(n+1)(n)(n-1)\cdots 3 \cdot 2 \cdot 1} = \frac{1}{n+1}, \\ \rho &= \lim_{n\to\infty} \rho_n = \lim_{n\to\infty} \frac{1}{n+1} = 0. \end{aligned}$$

Since $\rho < 1$, the series converges.

▶ **Example 2.** Test for convergence the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots.$$

We find

$$\begin{aligned} \rho_n &= \left| \frac{1}{n+1} \div \frac{1}{n} \right| = \frac{n}{n+1}, \\ \rho &= \lim_{n\to\infty} \frac{n}{n+1} = \lim_{n\to\infty} \frac{1}{1+\frac{1}{n}} = 1. \end{aligned}$$

Here the test tells us nothing and we must use some different test. A word of warning from this example: Notice that $\rho_n = n/(n+1)$ is always less than 1. Be careful not to confuse this ratio with ρ and conclude incorrectly that this series converges. (It is actually divergent as we proved by the integral test.) Remember that ρ is *not* the same as the ratio $\rho_n = |a_{n+1}/a_n|$, but is the *limit* of this ratio as $n \to \infty$.

▶ PROBLEMS, SECTION 6

Use the ratio test to find whether the following series converge or diverge:

18. $\displaystyle\sum_{n=1}^{\infty} \frac{2^n}{n^2}$ **19.** $\displaystyle\sum_{n=0}^{\infty} \frac{3^n}{2^{2n}}$ **20.** $\displaystyle\sum_{n=0}^{\infty} \frac{n!}{(2n)!}$

FEUE WHEP

21. $\displaystyle\sum_{n=0}^{\infty} \frac{5^n (n!)^2}{(2n)!}$ **22.** $\displaystyle\sum_{n=1}^{\infty} \frac{10^n}{(n!)^2}$ **23.** $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{100^n}$

24. $\displaystyle\sum_{n=0}^{\infty} \frac{3^{2n}}{2^{3n}}$ **25.** $\displaystyle\sum_{n=0}^{\infty} \frac{e^n}{\sqrt{n!}}$ **26.** $\displaystyle\sum_{n=0}^{\infty} \frac{(n!)^3 e^{3n}}{(3n)!}$

27. $\displaystyle\sum_{n=0}^{\infty} \frac{100^n}{n^{200}}$ **28.** $\displaystyle\sum_{n=0}^{\infty} \frac{n!(2n)!}{(3n)!}$ **29.** $\displaystyle\sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{n!}$

30. Prove the ratio test. *Hint:* If $|a_{n+1}/a_n| \to \rho < 1$, take σ so that $\rho < \sigma < 1$. Then $|a_{n+1}/a_n| < \sigma$ if n is large, say $n \geq N$. This means that we have $|a_{N+1}| < \sigma|a_N|$, $|a_{N+2}| < \sigma|a_{N+1}| < \sigma^2|a_N|$, and so on. Compare with the geometric series

$$\sum_{n=1}^{\infty} \sigma^n |a_N|.$$

Also prove that a series with $\rho > 1$ diverges. *Hint:* Take $\rho > \sigma > 1$, and use the preliminary test.

D. A Special Comparison Test

This test has two parts: (a) a convergence test, and (b) a divergence test. (See Problem 37.)

> (a) If $\sum_{n=1}^{\infty} b_n$ is a convergent series of positive terms and $a_n \geq 0$ and a_n/b_n tends to a (finite) limit, then $\sum_{n=1}^{\infty} a_n$ converges.
> (b) If $\sum_{n=1}^{\infty} d_n$ is a divergent series of positive terms and $a_n \geq 0$ and a_n/d_n tends to a limit greater than 0 (or tends to $+\infty$), then $\sum_{n=1}^{\infty} a_n$ diverges.

There are really two steps in using either of these tests, namely, to decide on a comparison series, and then to compute the required limit. The first part is the most important; given a good comparison series it is a routine process to find the needed limit. The method of finding the comparison series is best shown by examples.

Example 1. Test for convergence

$$\sum_{n=3}^{\infty} \frac{\sqrt{2n^2 - 5n + 1}}{4n^3 - 7n^2 + 2}.$$

Remember that whether a series converges or diverges depends on what the terms are as n becomes larger and larger. We are interested in the nth term as $n \to \infty$. Think of $n = 10^{10}$ or 10^{100}, say; a little calculation should convince you that as n increases, $2n^2 - 5n + 1$ is $2n^2$ to quite high accuracy. Similarly, the denominator in our example is nearly $4n^3$ for large n. By Section 9, fact 1, we see that the factor $\sqrt{2}/4$ in every term does not affect convergence. So we consider as a comparison series just

$$\sum_{n=3}^{\infty} \frac{\sqrt{n^2}}{n^3} = \sum_{n=3}^{\infty} \frac{1}{n^2}$$

10258

which we recognize (say by integral test) as a convergent series. Hence we use test (a) to try to show that the given series converges. We have:

$$\begin{aligned}\lim_{n\to\infty}\frac{a_n}{b_n} &= \lim_{n\to\infty}\left(\frac{\sqrt{2n^2-5n+1}}{4n^3-7n^2+2}\div\frac{1}{n^2}\right)\\ &= \lim_{n\to\infty}\frac{n^2\sqrt{2n^2-5n+1}}{4n^3-7n^2+2}\\ &= \lim_{n\to\infty}\frac{\sqrt{2-\frac{5}{n}+\frac{1}{n^2}}}{4-\frac{7}{n}+\frac{2}{n^3}} = \frac{\sqrt{2}}{4}.\end{aligned}$$

Since this is a finite limit, the given series converges. (With practice, you won't need to do all this algebra! You should be able to look at the original problem and see that, for large n, the terms are essentially $1/n^2$, so the series converges.)

▶ **Example 2.** Test for convergence

$$\sum_{n=2}^{\infty}\frac{3^n-n^3}{n^5-5n^2}.$$

Here we must first decide which is the important term as $n\to\infty$; is it 3^n or n^3? We can find out by comparing their logarithms since $\ln N$ and N increase or decrease together. We have $\ln 3^n = n\ln 3$, and $\ln n^3 = 3\ln n$. Now $\ln n$ is much smaller than n, so for large n we have $n\ln 3 > 3\ln n$, and $3^n > n^3$. (You might like to compute $100^3 = 10^6$, and $3^{100} > 5\times 10^{47}$.) The denominator of the given series is approximately n^5. Thus the comparison series is $\sum_{n=2}^{\infty} 3^n/n^5$. It is easy to prove this divergent by the ratio test. Now by test (b)

$$\lim_{n\to\infty}\left(\frac{3^n-n^3}{n^5-5n^2}\div\frac{3^n}{n^5}\right) = \lim_{n\to\infty}\frac{1-\frac{n^3}{3^n}}{1-\frac{5}{n^3}} = 1$$

which is greater than zero, so the given series diverges.

▶ PROBLEMS, SECTION 6

Use the special comparison test to find whether the following series converge or diverge.

31. $\displaystyle\sum_{n=9}^{\infty}\frac{(2n+1)(3n-5)}{\sqrt{n^2-73}}$

32. $\displaystyle\sum_{n=0}^{\infty}\frac{n(n+1)}{(n+2)^2(n+3)}$

33. $\displaystyle\sum_{n=5}^{\infty}\frac{1}{2^n-n^2}$

34. $\displaystyle\sum_{n=1}^{\infty}\frac{n^2+3n+4}{n^4+7n^3+6n-3}$

35. $\displaystyle\sum_{n=3}^{\infty}\frac{(n-\ln n)^2}{5n^4-3n^2+1}$

36. $\displaystyle\sum_{n=1}^{\infty}\frac{\sqrt{n^3+5n-1}}{n^2-\sin n^3}$

37. Prove the special comparison test. *Hint* (part a): If $a_n/b_n\to L$ and $M>L$, then $a_n < Mb_n$ for large n. Compare $\sum_{n=1}^{\infty} a_n$ with $\sum_{n=1}^{\infty} Mb_n$.

FEUE WHEP 10259

7. ALTERNATING SERIES

So far we have been talking about series of positive terms (including series of absolute values). Now we want to consider one important case of a series whose terms have mixed signs. An *alternating series* is a series whose terms are alternately plus and minus; for example,

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\cdots+\frac{(-1)^{n+1}}{n}+\cdots \tag{7.1}$$

is an alternating series. We ask two questions about an alternating series. Does it converge? Does it converge absolutely (that is, when we make all signs positive)? Let us consider the second question first. In this example the series of absolute values

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{n}+\cdots$$

is the harmonic series (6.1), which diverges. We say that the series (7.1) is not absolutely convergent. Next we must ask whether (7.1) converges as it stands. If it had turned out to be absolutely convergent, we would not have to ask this question since an absolutely convergent series is also convergent (Problem 9). However, a series which is not absolutely convergent may converge or it may diverge; we must test it further. For alternating series the test is very simple:

> **Test for alternating series.** An alternating series converges if the absolute value of the terms decreases steadily to zero, that is, if $|a_{n+1}| \le |a_n|$ and $\lim_{n\to\infty} a_n = 0$.

In our example $\dfrac{1}{n+1} < \dfrac{1}{n}$, and $\lim\limits_{n\to\infty} \dfrac{1}{n} = 0$, so (7.1) converges.

PROBLEMS, SECTION 7

Test the following series for convergence.

1. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$ **2.** $\displaystyle\sum_{n=1}^{\infty} \frac{(-2)^n}{n^2}$ **3.** $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$

4. $\displaystyle\sum_{n=1}^{\infty} \frac{(-3)^n}{n!}$ **5.** $\displaystyle\sum_{n=2}^{\infty} \frac{(-1)^n}{\ln n}$ **6.** $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n n}{n+5}$

7. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n n}{1+n^2}$ **8.** $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n \sqrt{10n}}{n+2}$

9. Prove that an absolutely convergent series $\sum_{n=1}^{\infty} a_n$ is convergent. *Hint:* Put $b_n = a_n + |a_n|$. Then the b_n are nonnegative; we have $|b_n| \le 2|a_n|$ and $a_n = b_n - |a_n|$.

10. The following alternating series are divergent (but you are not asked to prove this). Show that $a_n \to 0$. Why doesn't the alternating series test prove (incorrectly) that these series converge?

(a) $2-\frac{1}{2}+\frac{2}{3}-\frac{1}{4}+\frac{2}{5}-\frac{1}{6}+\frac{2}{7}-\frac{1}{8}\cdots$

(b) $\frac{1}{\sqrt{2}}-\frac{1}{2}+\frac{1}{\sqrt{3}}-\frac{1}{3}+\frac{1}{\sqrt{4}}-\frac{1}{4}+\frac{1}{\sqrt{5}}-\frac{1}{5}\cdots$

FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP

8. CONDITIONALLY CONVERGENT SERIES

A series like (7.1) which converges, but does not converge absolutely, is called *conditionally convergent*. You have to use special care in handling conditionally convergent series because the positive terms alone form a divergent series and so do the negative terms alone. If you rearrange the terms, you will probably change the sum of the series, and you may even make it diverge! It is possible to rearrange the terms to make the sum any number you wish. Let us do this with the alternating harmonic series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$. Suppose we want to make the sum equal to 1.5. First we take enough positive terms to add to just over 1.5. The first three positive terms do this:

$$1 + \frac{1}{3} + \frac{1}{5} = 1\frac{8}{15} > 1.5.$$

Then we take enough negative terms to bring the partial sum back under 1.5; the one term $-\frac{1}{2}$ does this. Again we add positive terms until we have a little more than 1.5, and so on. Since the terms of the series are decreasing in absolute value, we are able (as we continue this process) to get partial sums just a little more or a little less than 1.5 but always nearer and nearer to 1.5. But this is what convergence of the series to the sum 1.5 means: that the partial sums should approach 1.5. You should see that we could pick in advance *any* sum that we want, and rearrange the terms of this series to get it. Thus, we must not rearrange the terms of a conditionally convergent series since its convergence and its sum depend on the fact that the terms are added in a particular order.

Here is a physical example of such a series which emphasizes the care needed in applying mathematical approximations in physical problems. Coulomb's law in electricity says that the force between two charges is equal to the product of the charges divided by the square of the distance between them (in electrostatic units; to use other units, say SI, we need only multiply by a numerical constant). Suppose there are unit positive charges at $x = 0, \sqrt{2}, \sqrt{4}, \sqrt{6}, \sqrt{8}, \cdots$, and unit negative charges at $x = 1, \sqrt{3}, \sqrt{5}, \sqrt{7}, \cdots$. We want to know the total force acting on the unit positive charge at $x = 0$ due to all the other charges. The negative charges attract the charge at $x = 0$ and try to pull it to the right; we call the forces exerted by them positive, since they are in the direction of the positive x axis. The forces due to the positive charges are in the negative x direction, and we call them negative. For example, the force due to the positive charge at $x = \sqrt{2}$ is $-(1 \cdot 1)/\left(\sqrt{2}\right)^2 = -1/2$. The total force on the charge at $x = 0$ is, then,

$$F = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots. \tag{8.1}$$

Now we know that this series converges as it stands (Section 7). But we have also seen that its sum (even the fact that it converges) can be changed by rearranging the terms. Physically this means that the force on the charge at the origin depends not only on the size and position of the charges, but also on the *order* in which we place them in their positions! This may very well go strongly against your physical intuition. You feel that a physical problem like this should have a definite answer. Think of it this way. Suppose there are two crews of workers, one crew placing the positive charges and one placing the negative. If one crew works faster than the other, it is clear that the force at any stage may be far from the F of equation (8.1) because there are many extra charges of one sign. The crews can never place *all* the

FEUE WHEP
10261

charges because there are an infinite number of them. At any stage the forces which would arise from the positive charges that are not yet in place, form a divergent series; similarly, the forces due to the unplaced negative charges form a divergent series of the opposite sign. We cannot then stop at some point and say that the rest of the series is negligible as we could in the bouncing ball problem in Section 1. But if we specify the *order* in which the charges are to be placed, then the sum S of the series is determined (S is probably different from F in (8.1) unless the charges are placed alternately). Physically this means that the value of the force as the crews proceed comes closer and closer to S, and we can use the sum of the (properly arranged) *infinite* series as a good approximation to the force.

9. USEFUL FACTS ABOUT SERIES

We state the following facts for reference:

1. The convergence or divergence of a series is not affected by multiplying every term of the series by the same nonzero constant. Neither is it affected by changing a finite number of terms (for example, omitting the first few terms).

2. Two convergent series $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ may be added (or subtracted) term by term. (Adding "term by term" means that the nth term of the sum is $a_n + b_n$.) The resulting series is convergent, and its sum is obtained by adding (subtracting) the sums of the two given series.

3. The terms of an *absolutely convergent series* may be rearranged in any order without affecting either the convergence or the sum. This is *not true* of conditionally convergent series as we have seen in Section 8.

PROBLEMS, SECTION 9

Test the following series for convergence or divergence. Decide for yourself which test is easiest to use, but don't forget the preliminary test. Use the facts stated above when they apply.

1. $\displaystyle\sum_{n=1}^{\infty} \frac{n-1}{(n+2)(n+3)}$ **2.** $\displaystyle\sum_{n=1}^{\infty} \frac{n^2-1}{n^2+1}$ **3.** $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^{\ln 3}}$

4. $\displaystyle\sum_{n=0}^{\infty} \frac{n^2}{n^3+4}$ **5.** $\displaystyle\sum_{n=1}^{\infty} \frac{n}{n^3-4}$ **6.** $\displaystyle\sum_{n=0}^{\infty} \frac{(n!)^2}{(2n)!}$

7. $\displaystyle\sum_{n=0}^{\infty} \frac{(2n)!}{3^n(n!)^2}$ **8.** $\displaystyle\sum_{n=1}^{\infty} \frac{n^5}{5^n}$ **9.** $\displaystyle\sum_{n=1}^{\infty} \frac{n^n}{n!}$

10. $\displaystyle\sum_{n=2}^{\infty} (-1)^n \frac{n}{n-1}$ **11.** $\displaystyle\sum_{n=4}^{\infty} \frac{2n}{n^2-9}$ **12.** $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n^2-n}$

13. $\displaystyle\sum_{n=0}^{\infty} \frac{n}{(n^2+4)^{3/2}}$ **14.** $\displaystyle\sum_{n=2}^{\infty} \frac{(-1)^n}{n^2-n}$ **15.** $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n n!}{10^n}$

16. $\displaystyle\sum_{n=0}^{\infty} \frac{2+(-1)^n}{n^2+7}$ **17.** $\displaystyle\sum_{n=1}^{\infty} \frac{(n!)^3}{(3n)!}$ **18.** $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{2^{\ln n}}$

10262

19. $\frac{1}{2^2} - \frac{1}{3^2} + \frac{1}{2^3} - \frac{1}{3^3} + \frac{1}{2^4} - \frac{1}{3^4} + \cdots$

20. $\frac{1}{2} + \frac{1}{2^2} - \frac{1}{3} - \frac{1}{3^2} + \frac{1}{4} + \frac{1}{4^2} - \frac{1}{5} - \frac{1}{5^2} + \cdots$

21. $\sum_{n=1}^{\infty} a_n \quad \text{if } a_{n+1} = \frac{n}{2n+3} a_n$

22. (a) $\sum_{n=1}^{\infty} \frac{1}{3^{\ln n}}$ (b) $\sum_{n=1}^{\infty} \frac{1}{2^{\ln n}}$

(c) For what values of k is $\sum_{n=1}^{\infty} \frac{1}{k^{\ln n}}$ convergent?

▶ 10. POWER SERIES; INTERVAL OF CONVERGENCE

We have been discussing series whose terms were constants. Even more important and useful are series whose terms are functions of x. There are many such series, but in this chapter we shall consider series in which the nth term is a constant times x^n or a constant times $(x-a)^n$ where a is a constant. These are called *power series*, because the terms are multiples of powers of x or of $(x-a)$. In later chapters we shall consider Fourier series whose terms involve sines and cosines, and other series (Legendre, Bessel, etc.) in which the terms may be polynomials or other functions.

By definition, a power series is of the form

(10.1)
$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots \qquad \text{or}$$
$$\sum_{n=0}^{\infty} a_n (x-a)^n = a_0 + a_1(x-a) + a_2(x-a)^2 + a_3(x-a)^3 + \cdots,$$

where the coefficients a_n are constants. Here are some examples:

(10.2a) $$1 - \frac{x}{2} + \frac{x^2}{4} - \frac{x^3}{8} + \cdots + \frac{(-x)^n}{2^n} + \cdots,$$

(10.2b) $$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + \frac{(-1)^{n+1}x^n}{n} + \cdots,$$

(10.2c) $$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \frac{(-1)^{n+1}x^{2n-1}}{(2n-1)!} + \cdots,$$

(10.2d) $$1 + \frac{(x+2)}{\sqrt{2}} + \frac{(x+2)^2}{\sqrt{3}} + \cdots + \frac{(x+2)^n}{\sqrt{n+1}} + \cdots.$$

Whether a power series converges or not depends on the value of x we are considering. We often use the ratio test to find the values of x for which a series converges. We illustrate this by testing each of the four series (10.2). Recall that in the ratio test we divide term $n+1$ by term n and take the absolute value of this ratio to get ρ_n, and then take the limit of ρ_n as $n \to \infty$ to get ρ.

▶ **Example 1.** For (10.2a), we have

$$\rho_n = \left| \frac{(-x)^{n+1}}{2^{n+1}} \div \frac{(-x)^n}{2^n} \right| = \left| \frac{x}{2} \right|,$$
$$\rho = \left| \frac{x}{2} \right|.$$

10263

The series converges for $\rho < 1$, that is, for $|x/2| < 1$ or $|x| < 2$, and it diverges for $|x| > 2$ (see Problem 6.30). Graphically we consider the interval on the x axis between $x = -2$ and $x = 2$; for any x in this interval the series (10.2a) converges. The endpoints of the interval, $x = 2$ and $x = -2$, must be considered separately. When $x = 2$, (10.2a) is

$$1 - 1 + 1 - 1 + \cdots,$$

which is divergent; when $x = -2$, (10.2a) is $1 + 1 + 1 + 1 + \cdots$, which is divergent. Then the interval of convergence of (10.2a) is stated as $-2 < x < 2$.

Example 2. For (10.2b) we find

$$\rho_n = \left|\frac{x^{n+1}}{n+1} \div \frac{x^n}{n}\right| = \left|\frac{nx}{n+1}\right|,$$

$$\rho = \lim_{n\to\infty}\left|\frac{nx}{n+1}\right| = |x|.$$

The series converges for $|x| < 1$. Again we must consider the endpoints of the interval of convergence, $x = 1$ and $x = -1$. For $x = 1$, the series (10.2b) is $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$; this is the alternating harmonic series and is convergent. For $x = -1$, (10.2b) is $-1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{4} - \cdots$; this is the harmonic series (times -1) and is divergent. Then we state the interval of convergence of (10.2b) as $-1 < x \le 1$. Notice carefully how this differs from our result for (10.2a). Series (10.2a) did not converge at either endpoint and we used only $<$ signs in stating its interval of convergence. Series (10.2b) converges at $x = 1$, so we use the sign $\le$ to include $x = 1$. You must always test a series at its endpoints and include the results in your statement of the interval of convergence. A series may converge at neither, either one, or both of the endpoints.

Example 3. In (10.2c), the absolute value of the nth term is $|x^{2n-1}/(2n-1)!|$. To get term $n+1$ we replace n by $n+1$; then $2n-1$ is replaced by $2(n+1)-1 = 2n+1$, and the absolute value of term $n+1$ is

$$\left|\frac{x^{2n+1}}{(2n+1)!}\right|.$$

Thus we get

$$\rho_n = \left|\frac{x^{2n+1}}{(2n+1)!} \div \frac{x^{2n-1}}{(2n-1)!}\right| = \left|\frac{x^2}{(2n+1)(2n)}\right|,$$

$$\rho = \lim_{n\to\infty}\left|\frac{x^2}{(2n+1)(2n)}\right| = 0.$$

Since $\rho < 1$ for all values of x, this series converges for all x.

Example 4. In (10.2d), we find

$$\rho_n = \left|\frac{(x+2)^{n+1}}{\sqrt{n+2}} \div \frac{(x+2)^n}{\sqrt{n+1}}\right|,$$

$$\rho = \lim_{n\to\infty}\left|(x+2)\frac{\sqrt{n+1}}{\sqrt{n+2}}\right| = |x+2|.$$

FÉUE WHEP

The series converges for $|x+2|<1$; that is, for $-1<x+2<1$, or $-3<x<-1$. If $x=-3$, (10.2d) is

$$1-\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}-\frac{1}{\sqrt{4}}+\cdots$$

which is convergent by the alternating series test. For $x=-1$, the series is

$$1+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\cdots=\sum_{n=0}^{\infty}\frac{1}{\sqrt{n+1}}$$

which is divergent by the integral test. Thus, the series converges for $-3\le x<1$.

▶ PROBLEMS, SECTION 10

Find the interval of convergence of each of the following power series; be sure to investigate the endpoints of the interval in each case.

1. $\displaystyle\sum_{n=0}^{\infty}(-1)^n x^n$ **2.** $\displaystyle\sum_{n=0}^{\infty}\frac{(2x)^n}{3^n}$ **3.** $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n x^n}{n(n+1)}$

4. $\displaystyle\sum_{n=1}^{\infty}\frac{x^{2n}}{2^n n^2}$ **5.** $\displaystyle\sum_{n=1}^{\infty}\frac{x^n}{(n!)^2}$ **6.** $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n x^n}{(2n)!}$

7. $\displaystyle\sum_{n=1}^{\infty}\frac{x^{3n}}{n}$ **8.** $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n x^n}{\sqrt{n}}$ **9.** $\displaystyle\sum_{n=1}^{\infty}(-1)^n n^3 x^n$

10. $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n x^{2n}}{(2n)^{3/2}}$ **11.** $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{x}{5}\right)^n$ **12.** $\displaystyle\sum_{n=1}^{\infty}n(-2x)^n$

13. $\displaystyle\sum_{n=1}^{\infty}\frac{n(-x)^n}{n^2+1}$ **14.** $\displaystyle\sum_{n=1}^{\infty}\frac{n}{n+1}\left(\frac{x}{3}\right)^n$ **15.** $\displaystyle\sum_{n=1}^{\infty}\frac{(x-2)^n}{3^n}$

16. $\displaystyle\sum_{n=1}^{\infty}\frac{(x-1)^n}{2^n}$ **17.** $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n(x+1)^n}{n}$ **18.** $\displaystyle\sum_{n=1}^{\infty}\frac{(-2)^n(2x+1)^n}{n^2}$

The following series are *not* power series, but you can transform each one into a power series by a change of variable and so find out where it converges.

19. $\sum_0^\infty 8^{-n}(x^2-1)^n$ *Method:* Let $y=x^2-1$. The power series $\sum_0^\infty 8^{-n}y^n$ converges for $|y|<8$, so the original series converges for $|x^2-1|<8$, which means $|x|<3$.

20. $\displaystyle\sum_{0}^{\infty}(-1)^n\frac{2^n}{n!}(x^2+1)^{2n}$ **21.** $\displaystyle\sum_{2}^{\infty}\frac{(-1)^n x^{n/2}}{n\ln n}$

22. $\displaystyle\sum_{0}^{\infty}\frac{n!(-1)^n}{x^n}$ **23.** $\displaystyle\sum_{0}^{\infty}\frac{3^n(n+1)}{(x+1)^n}$

24. $\displaystyle\sum_{0}^{\infty}\left(\sqrt{x^2+1}\right)^n\frac{2^n}{3^n+n^3}$ **25.** $\displaystyle\sum_{0}^{\infty}(\sin x)^n(-1)^n 2^n$

FEUE WHEP
10265

11. THEOREMS ABOUT POWER SERIES

We have seen that a power series $\sum_{n=0}^{\infty} a_n x^n$ converges in some interval with center at the origin. For each value of x (in the interval of convergence) the series has a finite sum whose value depends, of course, on the value of x. Thus we can write the sum of the series as $S(x) = \sum_{n=0}^{\infty} a_n x^n$. We see then that a power series (within its interval of convergence) defines a function of x, namely $S(x)$. In describing the relation of the series and the function $S(x)$, we may say that the series converges to the function $S(x)$, or that the function $S(x)$ is represented by the series, or that the series is the power series of the function. Here we have thought of obtaining the function from a given series. We shall also (Section 12) be interested in finding a power series that converges to a given function. When we are working with power series and the functions they represent, it is useful to know the following theorems (which we state without proof; see advanced calculus texts). Power series are very useful and convenient because within their interval of convergence they can be handled much like polynomials.

1. A power series may be differentiated or integrated term by term; the resulting series converges to the derivative or integral of the function represented by the original series within the same interval of convergence as the original series (that is, not necessarily at the endpoints of the interval).

2. Two power series may be added, subtracted, or multiplied; the resultant series converges at least in the common interval of convergence. You may divide two series if the denominator series is not zero at $x = 0$, or if it is and the zero is canceled by the numerator [as, for example, in $(\sin x)/x$; see (13.1)]. The resulting series will have *some* interval of convergence (which can be found by the ratio test or more simply by complex variable theory—see Chapter 2, Section 7).

3. One series may be substituted in another provided that the values of the substituted series are in the interval of convergence of the other series.

4. The power series of a function is unique, that is, there is just one power series of the form $\sum_{n=0}^{\infty} a_n x^n$ which converges to a given function.

12. EXPANDING FUNCTIONS IN POWER SERIES

Very often in applied work, it is useful to find power series that represent given functions. We illustrate one method of obtaining such series by finding the series for $\sin x$. In this method we *assume* that there *is* such a series (see Section 14 for discussion of this point) and set out to find what the coefficients in the series must be. Thus we write

$$\sin x = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots \tag{12.1}$$

and try to find numerical values of the coefficients a_n to make (12.1) an identity (within the interval of convergence of the series). Since the interval of convergence of a power series contains the origin, (12.1) must hold when $x = 0$. If we substitute $x = 0$ into (12.1), we get $0 = a_0$ since $\sin 0 = 0$ and all the terms except a_0 on the

FEUE WHEP

right-hand side of the equation contain the factor x. Then to make (12.1) valid at $x = 0$, we must have $a_0 = 0$. Next we differentiate (12.1) term by term to get

$$\cos x = a_1 + 2a_2x + 3a_3x^2 + \cdots . \tag{12.2}$$

(This is justified by Theorem 1 of Section 11.) Again putting $x = 0$, we get $1 = a_1$. We differentiate again, and put $x = 0$ to get

$$\begin{aligned} -\sin x &= 2a_2 + 3\cdot 2a_3x + 4\cdot 3a_4x^2 + \cdots , \\ 0 &= 2a_2. \end{aligned} \tag{12.3}$$

Continuing the process of taking successive derivatives of (12.1) and putting $x = 0$, we get

$$\begin{aligned} -\cos x &= 3\cdot 2a_3 + 4\cdot 3\cdot 2a_4x + \cdots , \\ -1 &= 3!\,a_3, \qquad a_3 = -\frac{1}{3!}; \\ \sin x &= 4\cdot 3\cdot 2\cdot a_4 + 5\cdot 4\cdot 3\cdot 2a_5x + \cdots , \\ 0 &= a_4; \\ \cos x &= 5\cdot 4\cdot 3\cdot 2a_5 + \cdots , \\ 1 &= 5!\,a_5, \cdots . \end{aligned} \tag{12.4}$$

We substitute these values back into (12.1) and get

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots . \tag{12.5}$$

You can probably see how to write more terms of this series without further computation. The $\sin x$ series converges for all x; see Example 3, Section 10.

Series obtained in this way are called *Maclaurin series* or *Taylor series about the origin*. A Taylor series in general means a series of powers of $(x-a)$, where a is some constant. It is found by writing $(x-a)$ instead of x on the right-hand side of an equation like (12.1), differentiating just as we have done, but substituting $x = a$ instead of $x = 0$ at each step. Let us carry out this process in general for a function $f(x)$. As above, we assume that there is a Taylor series for $f(x)$, and write

$$\begin{aligned} f(x) &= a_0 + a_1(x-a) + a_2(x-a)^2 + a_3(x-a)^3 + a_4(x-a)^4 + \cdots \\ &\quad + a_n(x-a)^n + \cdots , \\ f'(x) &= a_1 + 2a_2(x-a) + 3a_3(x-a)^2 + 4a_4(x-a)^3 + \cdots \\ &\quad + na_n(x-a)^{n-1} + \cdots , \\ f''(x) &= 2a_2 + 3\cdot 2a_3(x-a) + 4\cdot 3a_4(x-a)^2 + \cdots \\ &\quad + n(n-1)a_n(x-a)^{n-2} + \cdots , \\ f'''(x) &= 3!\,a_3 + 4\cdot 3\cdot 2a_4(x-a) + \cdots \\ &\quad + n(n-1)(n-2)a_n(x-a)^{n-3} + \cdots , \\ &\ \ \vdots \\ f^{(n)}(x) &= n(n-1)(n-2)\cdots 1\cdot a_n + \text{terms containing powers of } (x-a). \end{aligned} \tag{12.6}$$

10267

[The symbol $f^{(n)}(x)$ means the nth derivative of $f(x)$.] We now put $x = a$ in each equation of (12.6) and obtain

$$\begin{aligned} f(a) = a_0, \quad f'(a) = a_1, \quad f''(a) = 2a_2, \\ f'''(a) = 3!\, a_3, \quad \cdots, \quad f^{(n)}(a) = n!\, a_n. \end{aligned} \tag{12.7}$$

[Remember that $f'(a)$ means to differentiate $f(x)$ and then put $x = a$; $f''(a)$ means to find $f''(x)$ and then put $x = a$, and so on.]

We can then write the Taylor series for $f(x)$ about $x = a$:

$$f(x) = f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \cdots + \frac{1}{n!}(x-a)^n f^{(n)}(a) + \cdots. \tag{12.8}$$

The Maclaurin series for $f(x)$ is the Taylor series about the origin. Putting $a = 0$ in (12.8), we obtain the Maclaurin series for $f(x)$:

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \frac{x^3}{3!}f'''(0) + \cdots + \frac{x^n}{n!}f^{(n)}(0) + \cdots. \tag{12.9}$$

We have written this in general because it is sometimes convenient to have the formulas for the coefficients. However, finding the higher order derivatives in (12.9) for any but the simplest functions is unnecessarily complicated (try it for, say, $e^{\tan x}$). In Section 13, we shall discuss much easier ways of getting Maclaurin and Taylor series by combining a few basic series. Meanwhile, you should verify (Problem 1, below) the basic series (13.1) to (13.5) and memorize them.

PROBLEMS, SECTION 12

1. By the method used to obtain (12.5) [which is the series (13.1) below], verify each of the other series (13.2) to (13.5) below.

13. TECHNIQUES FOR OBTAINING POWER SERIES EXPANSIONS

There are often simpler ways for finding the power series of a function than the successive differentiation process in Section 12. Theorem 4 in Section 11 tells us that for a given function there is *just one* power series, that is, just one series of the form $\sum_{n=0}^{\infty} a_n x^n$. Therefore we can obtain it by any correct method and be sure that it is the same Maclaurin series we would get by using the method of Section 12. We shall illustrate a variety of methods for obtaining power series. First of all, it is a great timesaver for you to verify (Problem 12.1) and then memorize the basic series (13.1) to (13.5). We shall use these series without further derivation when we need them.

FEUE WHEP

convergent for

(13.1) $$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots, \qquad \text{all } x;$$

(13.2) $$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots, \qquad \text{all } x;$$

(13.3) $$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots, \qquad \text{all } x;$$

(13.4)
$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots, \qquad -1 < x \le 1;$$

(13.5) $$(1+x)^p = \sum_{n=0}^{\infty} \binom{p}{n} x^n = 1 + px + \frac{p(p-1)}{2!}x^2 + \frac{p(p-1)(p-2)}{3!}x^3 + \cdots, \qquad |x| < 1,$$

(binomial series; p is any real number, positive or negative and $\binom{p}{n}$ is called a binomial coefficient—see method C below.)

When we use a series to approximate a function, we may want only the first few terms, but in derivations, we may want the formula for the general term so that we can write the series in summation form. Let's look at some methods of obtaining either or both of these results.

A. Multiplying a Series by a Polynomial or by Another Series

▶ **Example 1.** To find the series for $(x+1)\sin x$, we multiply $(x+1)$ times the series (13.1) and collect terms:

$$(x+1)\sin x = (x+1)\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right)$$
$$= x + x^2 - \frac{x^3}{3!} - \frac{x^4}{3!} + \cdots.$$

You can see that this is easier to do than taking the successive derivatives of the product $(x+1)\sin x$, and Theorem 4 assures us that the results are the same.

10269

Example 2. To find the series for $e^x \cos x$, we multiply (13.2) by (13.3):

$$\begin{aligned}
e^x \cos x &= \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\right)\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots\right) \\
&= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} \cdots \\
&\qquad\quad - \frac{x^2}{2!} - \frac{x^3}{2!} - \frac{x^4}{2!\,2!} \cdots \\
&\qquad\qquad\qquad\qquad + \frac{x^4}{4!} \cdots \\
\hline
&= 1 + x + 0x^2 - \frac{x^3}{3} - \frac{x^4}{6} \cdots = 1 + x - \frac{x^3}{3} - \frac{x^4}{6} \cdots .
\end{aligned}$$

There are two points to note here. First, as you multiply, line up the terms involving each power of x in a column; this makes it easier to combine them. Second, be careful to include *all* the terms in the product out to the power you intend to stop with, but don't include *any* higher powers. In the above example, note that we did not include the $x^3 \cdot x^2$ terms; if we wanted the x^5 term in the answer, we would have to include *all* products giving x^5 (namely, $x \cdot x^4, x^3 \cdot x^2$, and $x^5 \cdot 1$).

Also see Chapter 2, Problem 17.30, for a simple way of getting the general term of this series.

B. Division of Two Series or of a Series by a Polynomial

Example 1. To find the series for $(1/x)\ln(1+x)$, we divide (13.4) by x. You should be able to do this in your head and just write down the answer.

$$\frac{1}{x}\ln(1+x) = 1 - \frac{x}{2} + \frac{x^2}{3} - \frac{x^3}{4} + \cdots .$$

To obtain the summation form, we again just divide (13.4) by x. We can simplify the result by changing the limits to start at $n = 0$, that is, replace n by $n + 1$.

$$\frac{1}{x}\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^{n-1}}{n} = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n+1}.$$

FEUE WHEP

▶ **Example 2.** To find the series for $\tan x$, we divide the series for $\sin x$ by the series for $\cos x$ by long division:

$$
\begin{array}{r|l}
 & x + \dfrac{x^3}{3} + \dfrac{2}{15}x^5 \cdots \\ \hline
1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} \cdots & x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} \cdots \\
 & x - \dfrac{x^3}{2!} + \dfrac{x^5}{4!} \cdots \\ \cline{2-2}
 & \qquad \dfrac{x^3}{3} - \dfrac{x^5}{30} \cdots \\
 & \qquad \dfrac{x^3}{3} - \dfrac{x^5}{6} \cdots \\ \cline{2-2}
 & \qquad\qquad \dfrac{2x^5}{15} \cdots \text{ , etc.}
\end{array}
$$

C. Binomial Series

If you recall the binomial theorem, you may see that (13.5) looks just like the beginning of the binomial theorem for the expansion of $(a+b)^n$ if we put $a = 1$, $b = x$, and $n = p$. The difference here is that we allow p to be negative or fractional, and in these cases the expansion is an infinite series. The series converges for $|x| < 1$ as you can verify by the ratio test. (See Problem 1.)

From (13.5), we see that the binomial coefficients are:

$$
\begin{aligned}
\binom{p}{0} &= 1, \\
\binom{p}{1} &= p, \\
\binom{p}{2} &= \frac{p(p-1)}{2!}, \\
\binom{p}{3} &= \frac{p(p-1)(p-2)}{3!}, \cdots, \\
\binom{p}{n} &= \frac{p(p-1)(p-2)\cdots(p-n+1)}{n!}.
\end{aligned}
\tag{13.6}
$$

▶ **Example 1.** To find the series for $1/(1+x)$, we use the binomial series (13.5) to write

$$
\begin{aligned}
\frac{1}{1+x} = (1+x)^{-1} &= 1 - x + \frac{(-1)(-2)}{2!}x^2 + \frac{(-1)(-2)(-3)}{3!}x^3 + \cdots \\
&= 1 - x + x^2 - x^3 + \cdots = \sum_{n=0}^{\infty}(-x)^n.
\end{aligned}
$$

FEUE WHEP

Example 2. The series for $\sqrt{1+x}$ is (13.5) with $p = 1/2$.

$$\begin{aligned}\sqrt{1+x} = (1+x)^{1/2} &= \sum_{n=0}^{\infty}\binom{1/2}{n}x^n\\ &= 1 + \frac{1}{2}x + \frac{\frac{1}{2}(-\frac{1}{2})}{2!}x^2 + \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})}{3!}x^3 + \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})(-\frac{5}{2})}{4!}x^4 + \cdots\\ &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 \cdots .\end{aligned}$$

From (13.6) we can see that the binomial coefficients when $n = 0$ and $n = 1$ are $\binom{1/2}{0} = 1$ and $\binom{1/2}{1} = 1/2$. For $n \geq 2$, we can write

$$\begin{aligned}\binom{\frac{1}{2}}{n} &= \frac{(\frac{1}{2})(-\frac{1}{2})(-\frac{3}{2})\cdots(\frac{1}{2}-n+1)}{n!} = \frac{(-1)^{n-1}3\cdot 5\cdot 7\cdots(2n-3)}{n!\,2^n}\\ &= \frac{(-1)^{n-1}(2n-3)!!}{(2n)!!}\end{aligned}$$

where the double factorial of an odd number means the product of that number times all smaller odd numbers, and a similar definition for even numbers. For example, $7!! = 7\cdot 5\cdot 3$, and $8!! = 8\cdot 6\cdot 4\cdot 2$.

PROBLEMS, SECTION 13

1. Use the ratio test to show that a binomial series converges for $|x| < 1$.
2. Show that the binomial coefficients $\binom{-1}{n} = (-1)^n$.
3. Show that if p is a positive integer, then $\binom{p}{n} = 0$ when $n > p$, so $(1+x)^p = \sum \binom{p}{n}x^n$ is just a sum of $p+1$ terms, from $n = 0$ to $n = p$. For example, $(1+x)^2$ has 3 terms, $(1+x)^3$ has 4 terms, etc. This is just the familiar binomial theorem.
4. Write the Maclaurin series for $1/\sqrt{1+x}$ in $\sum$ form using the binomial coefficient notation. Then find a formula for the binomial coefficients in terms of n as we did in Example 2 above.

D. Substitution of a Polynomial or a Series for the Variable in Another Series

Example 1. Find the series for e^{-x^2}. Since we know the series (13.3) for e^x, we simply replace the x there by $-x^2$ to get

$$\begin{aligned}e^{-x^2} &= 1 - x^2 + \frac{(-x^2)^2}{2!} + \frac{(-x^2)^3}{3!} + \cdots\\ &= 1 - x^2 + \frac{(x^4)}{2!} - \frac{x^6}{3!} + \cdots .\end{aligned}$$

Example 2. Find the series for $e^{\tan x}$. Here we must replace the x in (13.3) by the series of Example 2 in method B. Let us agree in advance to keep terms only as far as x^4; we then write only terms which can give rise to powers of x up to 4, and neglect

10272

any higher powers:

$$\begin{aligned}
e^{\tan x} &= 1 + \left(x + \frac{x^3}{3} + \cdots\right) + \frac{1}{2!}\left(x + \frac{x^3}{3} + \cdots\right)^2 \\
&\quad + \frac{1}{3!}\left(x + \frac{x^3}{3} + \cdots\right)^3 + \frac{1}{4!}(x + \cdots)^4 + \cdots \\
&= 1 + x \qquad + \frac{x^3}{3} \qquad\qquad + \cdots \\
&\qquad + \frac{x^2}{2!} \qquad + \frac{2x^4}{3 \cdot 2!} + \cdots \\
&\qquad\qquad + \frac{x^3}{3!} + \frac{x^4}{4!} \quad + \cdots \\
\hline
&= 1 + x + \frac{x^2}{2} + \frac{x^3}{2} + \frac{3}{8}x^4 \quad + \cdots .
\end{aligned}$$

E. Combination of Methods

▶ **Example.** Find the series for $\operatorname{arc}\tan x$. Since

$$\int_0^x \frac{dt}{1+t^2} = \operatorname{arc}\tan t \Big|_0^x = \operatorname{arc}\tan x,$$

we first write out (as a binomial series) $(1+t^2)^{-1}$ and then integrate term by term:

$$\begin{aligned}
&(1+t^2)^{-1} = 1 - t^2 + t^4 - t^6 + \cdots ; \\
&\int_0^x \frac{dt}{1+t^2} = t - \frac{t^3}{3} + \frac{t^5}{5} - \frac{t^7}{7} + \cdots \Big|_0^x .
\end{aligned}$$

Thus, we have

$$\operatorname{arc}\tan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots . \tag{13.7}$$

Compare this simple way of getting the series with the method in Section 12 of finding successive derivatives of $\operatorname{arc}\tan x$.

F. Taylor Series Using the Basic Maclaurin Series

In many simple cases it is possible to obtain a Taylor series using the basic memorized Maclaurin series instead of the formulas or method of Section 12.

▶ **Example 1.** Find the first few terms of the Taylor series for $\ln x$ about $x = 1$. [This means a series of powers of $(x-1)$ rather than powers of x.] We write

$$\ln x = \ln[1 + (x-1)]$$

and use (13.4) with x replaced by $(x-1)$:

$$\ln x = \ln[1 + (x-1)] = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 \cdots .$$

10273

Example 2. Expand $\cos x$ about $x = 3\pi/2$. We write

$$\cos x = \cos\left[\frac{3\pi}{2} + \left(x - \frac{3\pi}{2}\right)\right] = \sin\left(x - \frac{3\pi}{2}\right)$$
$$= \left(x - \frac{3\pi}{2}\right) - \frac{1}{3!}\left(x - \frac{3\pi}{2}\right)^3 + \frac{1}{5!}\left(x - \frac{3\pi}{2}\right)^5 \cdots$$

using (13.1) with x replaced by $(x - 3\pi/2)$.

G. Using a Computer

You can also do problems like these using a computer. This is a good method for complicated functions where it saves you a lot of algebra. However, you're not saving time if it takes longer to type a problem into the computer than to do it in your head! For example, you should be able to just write down the first few terms of $(\sin x)/x$ or $(1 - \cos x)/x^2$. A good method of study is to practice doing problems by hand and also check your results using the computer. This will turn up errors you are making by hand, and also let you discover what the computer will do and what it won't do! It is very illuminating to computer plot the function you are expanding, along with several partial sums of the series, in order to see how accurately the partial sums represent the function—see the following example.

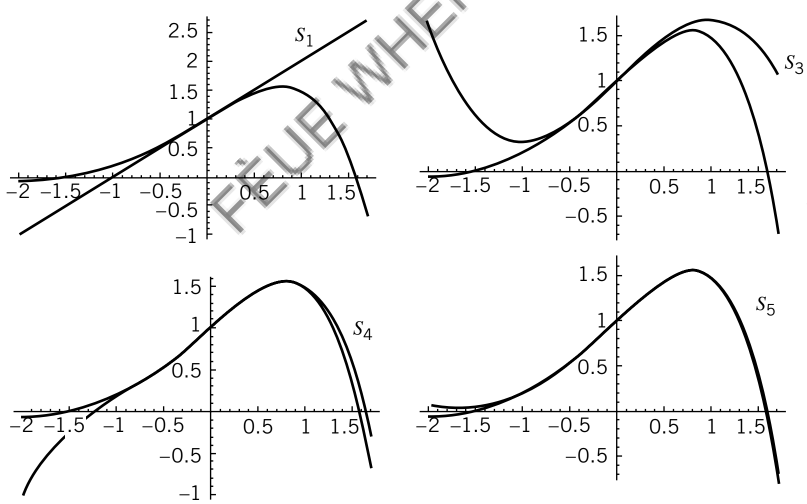


Figure 13.1

Example. Plot the function $e^x \cos x$ together with several partial sums of its Maclaurin series. Using Example 2 in 13A or a computer, we have

$$e^x \cos x = 1 + x - \frac{x^3}{3} - \frac{x^4}{6} - \frac{x^5}{30} \cdots.$$

Figure 13.1 shows plots of the function along with each of the partial sums $S_1 = 1+x$, $S_3 = 1 + x - \frac{x^3}{3}$, $S_4 = 1 + x - \frac{x^3}{3} - \frac{x^4}{6}$, $S_5 = 1 + x - \frac{x^3}{3} - \frac{x^4}{6} - \frac{x^5}{30}$. We can see from the graphs the values of x for which an approximation is fairly good. Also see Section 14.

▶ PROBLEMS, SECTION 13

Using the methods of this section:

(a) Find the first few terms of the Maclaurin series for each of the following functions.

(b) Find the general term and write the series in summation form.

(c) Check your results in (a) by computer.

(d) Use a computer to plot the function and several approximating partial sums of the series.

5. $x^2\ln(1-x)$ **6.** $x\sqrt{1+x}$ **7.** $\dfrac{1}{x}\sin x$

8. $\dfrac{1}{\sqrt{1-x^2}}$ **9.** $\dfrac{1+x}{1-x}$ **10.** $\sin x^2$

11. $\dfrac{\sin\sqrt{x}}{\sqrt{x}},\quad x>0$ **12.** $\displaystyle\int_0^x \cos t^2\,dt$ **13.** $\displaystyle\int_0^x e^{-t^2}\,dt$

14. $\ln\sqrt{\dfrac{1+x}{1-x}}=\displaystyle\int_0^x \frac{dt}{1-t^2}$ **15.** $\arcsin x=\displaystyle\int_0^x \frac{dt}{\sqrt{1-t^2}}$

16. $\cosh x=\dfrac{e^x+e^{-x}}{2}$ **17.** $\ln\dfrac{1+x}{1-x}$

18. $\displaystyle\int_0^x \frac{\sin t\,dt}{t}$ **19.** $\ln(x+\sqrt{1+x^2}\,)=\displaystyle\int_0^x \frac{dt}{\sqrt{1+t^2}}$

Find the first few terms of the Maclaurin series for each of the following functions and check your results by computer.

20. $e^x\sin x$ **21.** $\tan^2 x$ **22.** $\dfrac{e^x}{1-x}$

23. $\dfrac{1}{1+x+x^2}$ **24.** $\sec x=\dfrac{1}{\cos x}$ **25.** $\dfrac{2x}{e^{2x}-1}$

26. $\dfrac{1}{\sqrt{\cos x}}$ **27.** $e^{\sin x}$ **28.** $\sin[\ln(1+x)]$

29. $\sqrt{1+\ln(1+x)}$ **30.** $\sqrt{\dfrac{1-x}{1+x}}$ **31.** $\cos(e^x-1)$

32. $\ln(1+xe^x)$ **33.** $\dfrac{1-\sin x}{1-x}$ **34.** $\ln(2-e^{-x})$

35. $\dfrac{x}{\sin x}$ **36.** $\displaystyle\int_0^u \frac{\sin x\,dx}{\sqrt{1-x^2}}$

37. $\ln\cos x$ *Hints:* Method 1: Write $\cos x=1+(\cos x-1)=1+u$; use the series you know for $\ln(1+u)$; replace u by the Maclaurin series for $(\cos x-1)$. Method 2: $\ln\cos x=-\int_0^x \tan u\,du$. Use the series of Example 2 in method B.

38. $e^{\cos x}$ *Hint:* $e^{\cos x}=e\cdot e^{\cos x-1}$.

Using method F above, find the first few terms of the Taylor series for the following functions about the given points.

39. $f(x)=\sin x,\quad a=\pi/2$ **40.** $f(x)=\dfrac{1}{x},\quad a=1$

41. $f(x)=e^x,\quad a=3$ **42.** $f(x)=\cos x,\quad a=\pi$

43. $f(x)=\cot x,\quad a=\pi/2$ **44.** $f(x)=\sqrt{x}\quad a=25$

FEUE WHEP
10275

14. ACCURACY OF SERIES APPROXIMATIONS

The thoughtful student might well be disturbed about the mathematical manipulations we have been doing. How do we know whether these processes we have shown really give us series that approximate the functions being expanded? Certainly *some* functions cannot be expanded in a power series; since a power series becomes just a_0 when $x = 0$, it cannot be equal to any function (like $1/x$ or $\ln x$) which is infinite at the origin. So we might ask whether there are other functions (besides those that become infinite at the origin) which cannot be expanded in a power series. All we have done so far is to show methods of finding the power series for a function *if it has one*. Now is there a chance that there might be some functions which do not have series expansions, but for which our formal methods would give us a spurious series? Unfortunately, the answer is "Yes"; fortunately, this is not a very common difficulty in practice. However, you should know of the possibility and what to do about it. You may first think of the fact that, say, the equation

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots$$

is not valid for $|x| \geq 1$. This is a fairly easy restriction to determine; from the beginning we recognized that we could use our series expansions only when they converged. But there is another difficulty which can arise. It is possible for a series found by the above methods to converge and still not represent the function being expanded! A simple example of this is $e^{-(1/x^2)}$ for which the formal series is $0+0+0+\cdots$ because $e^{-(1/x^2)}$ and all its derivatives are zero at the origin (Problem 15.26). It is clear that $e^{-(1/x^2)}$ is not zero for $x^2 > 0$, so the series is certainly not correct. You can startle your friends with the following physical interpretation of this. Suppose that at $t = 0$ a car is at rest (zero velocity), and has zero acceleration, zero rate of change of acceleration, etc. (all derivatives of distance with respect to time are zero at $t = 0$). Then according to Newton's second law (force equals mass times acceleration), the instantaneous force acting on the car is also zero (and, in fact, so are all the derivatives of the force). Now we ask "Is it possible for the car to be moving immediately after $t = 0$?" The answer is "Yes"! For example, let its distance from the origin as a function of time be $e^{-(1/t^2)}$.

This strange behavior is really the fault of the function itself and not of our method of finding series. The most satisfactory way of avoiding the difficulty is to recognize, by complex variable theory, when functions can or cannot have power series. We shall consider this in Chapter 14, Section 2. Meanwhile, let us consider two important questions: (1) Does the Taylor or Maclaurin series in (12.8) or (12.9) actually converge to the function being expanded? (2) In a computation problem, if we know that a series converges to a given function, how rapidly does it converge? That is, how many terms must we use to get the accuracy we require? We take up these questions in order.

The *remainder* $R_n(x)$ in a Taylor series is the difference between the value of the function and the sum of $n + 1$ terms of the series:

$$R_n(x) = f(x) - \left[f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \cdots + \frac{1}{n!}(x-a)^n f^{(n)}(a)\right]. \tag{14.1}$$

FEUE WHEP

Saying that the series converges to the function means that $\lim_{n\to\infty} |R_n(x)| = 0$. There are many different formulas for $R_n(x)$ which are useful for special purposes; you can find these in calculus books. One such formula is

$$R_n(x) = \frac{(x-a)^{n+1} f^{(n+1)}(c)}{(n+1)!} \tag{14.2}$$

where c is some point between a and x. You can use this formula in some simple cases to prove that the Taylor or Maclaurin series for a function does converge to the function (Problems 11 to 13).

Error in Series Approximations Now suppose that we know in advance that the power series of a function does converge to the function (within the interval of convergence), and we want to use a series approximation for the function. We would like to estimate the error caused by using only a few terms of the series.

There is an easy way to estimate this error when the series is alternating and meets the alternating series test for convergence (Section 7). In this case the error is (in absolute value) less than the absolute value of the first neglected term (see Problem 1).

$$\text{If } S = \sum_{n=1}^{\infty} a_n \text{ is an alternating series with } |a_{n+1}| < |a_n|, \\ \text{and } \lim_{n\to\infty} a_n = 0, \text{ then } |S - (a_1 + a_2 + \cdots + a_n)| \le |a_{n+1}|. \tag{14.3}$$

▶ **Example 1.** Consider the series

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \frac{1}{32} + \frac{1}{64} \cdots .$$

The sum of this series [see (1.8), $a = 1, r = -\frac{1}{2}$] is $S = \frac{2}{3} = 0.666\cdots$. The sum of the terms through $-\frac{1}{32}$ is 0.656+, which differs from S by about 0.01. This is less than the next term $= \frac{1}{64} = 0.015+$.

Estimating the error by the first neglected term may be quite misleading for convergent series that are not alternating.

▶ **Example 2.** Suppose we approximate $\sum_{n=1}^{\infty} 1/n^2$ by the sum of the first five terms; the error is then about 0.18 [see problem 2(a)]. But the first neglected term is $1/6^2 = 0.028$ which is much less than the error. However, note that we are finding the sum of the power series $\sum_{n=1}^{\infty} x^n/n^2$ when $x = 1$, which is the largest x for which the series converges. If, instead, we ask for the sum of the series when $x = 1/2$, we find [see Problem 2(b)]:

$$S = \sum_{n=1}^{\infty} \frac{1}{n^2}\left(\frac{1}{2}\right)^n = 0.5822+ .$$

The sum of the first five terms of the series is 0.5815+, so the error is about 0.0007. The next term is $(\frac{1}{6})^2/6^2 = 0.0004$, which is less than the error but still of the

FÊUE WHEP

same order of magnitude. We can state the following theorem [Problem 2(c)] which covers many practical problems.

$$\text{If } S = \sum_{n=0}^{\infty} a_n x^n \text{ converges for } |x| < 1, \text{ and if}$$

$$|a_{n+1}| < |a_n| \text{ for } n > N, \text{ then} \tag{14.4}$$

$$\left| S - \sum_{n=0}^{N} a_n x^n \right| < |a_{N+1} x^{N+1}| \div (1 - |x|).$$

That is, as in (14.3), the error may be estimated by the first neglected term, but here the error may be a few times as large as the first neglected term instead of smaller. In the example of $\sum x^n/n^2$ with $x = \frac{1}{2}$, we have $1 - x = \frac{1}{2}$, so (14.4) says that the error is less than two times the next term. We observe that the error 0.0007 is less than 2(0.0004) as (14.4) says.

For values of $|x|$ much less than 1, $1 - |x|$ is about 1, so the next term gives a good error estimate in this case. If the interval of convergence is not $|x| < 1$, but, for example, $|x| < 2$ as in

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \left(\frac{x}{2}\right)^n,$$

we can easily let $x/2 = y$, and apply the theorem in terms of y.

PROBLEMS, SECTION 14

1. Prove theorem (14.3). *Hint:* Group the terms in the error as $(a_{n+1} + a_{n+2}) + (a_{n+3} + a_{n+4}) + \cdots$ to show that the error has the same sign as a_{n+1}. Then group them as $a_{n+1} + (a_{n+2} + a_{n+3}) + (a_{n+4} + a_{n+5}) + \cdots$ to show that the error has magnitude less than $|a_{n+1}|$.

2. (a) Using computer or tables (or see Chapter 7, Section 11), verify that $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6 = 1.6449+$, and also verify that the error in approximating the sum of the series by the first five terms is approximately 0.1813.

(b) By computer or tables verify that $\sum_{n=1}^{\infty}(1/n^2)(1/2)^n = \pi^2/12 - (1/2)(\ln 2)^2 = 0.5822+$, and that the sum of the first five terms is 0.5815+.

(c) Prove theorem (14.4). *Hint:* The error is $|\sum_{N+1}^{\infty} a_n x^n|$. Use the fact that the absolute value of a sum is less than or equal to the sum of the absolute values. Then use the fact that $|a_{n+1}| \le |a_n|$ to replace all a_n by a_{N+1}, and write the appropriate inequality. Sum the geometric series to get the result.

In Problems 3 to 7, assume that the Maclaurin series converges to the function.

3. If $0 < x < \frac{1}{2}$, show [using theorem (14.3)] that $\sqrt{1+x} = 1 + \frac{1}{2}x$ with an error less than 0.032. *Hint:* Note that the series is alternating after the first term.

4. Show that $\sin x = x$ with an error less than 0.021 for $0 < x < \frac{1}{2}$, and with an error less than 0.0002 for $0 < x < 0.1$. *Hint:* Use theorem (14.3) and note that the "next" term is the x^3 term.

5. Show that $1 - \cos x = x^2/2$ with an error less than 0.003 for $|x| < \frac{1}{2}$.

FEUE WHEP
10278

6. Show that $\ln(1-x) = -x$ with an error less than 0.0056 for $|x| < 0.1$. *Hint:* Use theorem (14.4).

7. Show that $2/\sqrt{4-x} = 1 + \frac{1}{8}x$ with an error less than $\frac{1}{32}$ for $0 < x < 1$. *Hint:* Let $x = 4y$, and use theorem (14.4).

8. Estimate the error if $\sum_{n=1}^{\infty} x^n/n^3$ is approximated by the sum of its first three terms for $|x| < \frac{1}{2}$.

9. Consider the series in Problem 4.6 and show that the remainder after n terms is $R_n = 1/(n+1)$. Compare the value of term $n+1$ with R_n for $n = 3, n = 10, n = 100, n = 500$ to see that the first neglected term is not a useful estimate of the error.

10. Show that the interval of convergence of the series $\sum_{n=1}^{\infty} x^n/(n^2+n)$ is $|x| \le 1$. (For $x = 1$, this is the series of Problem 9.) Using theorem (14.4), show that for $x = \frac{1}{2}$, four terms will give two decimal place accuracy.

11. Show that the Maclaurin series for $\sin x$ converges to $\sin x$. *Hint:* If $f(x) = \sin x$, $f^{(n+1)}(x) = \pm \sin x$ or $\pm \cos x$, and so $|f^{(n+1)}(x)| \le 1$ for all x and all n. Let $n \to \infty$ in (14.2).

12. Show as in Problem 11 that the Maclaurin series for e^x converges to e^x.

13. Show that the Maclaurin series for $(1+x)^p$ converges to $(1+x)^p$ when $0 < x < 1$.

▶ 15. SOME USES OF SERIES

In this chapter we are going to consider a few rather straightforward uses of series. In later chapters there will also be many other cases where we need them.

Numerical Computation With computers and calculators so available, you may wonder why we would ever want to use series for numerical computation. Here is an example to warn you of the pitfalls of blind computation.

▶ **Example 1.** Evaluate $f(x) = \ln\sqrt{(1+x)/(1-x)} - \tan x$ at $x = 0.0015$.

Here are answers from several calculators and computers: -9×10^{-16}, 3×10^{-10}, 6.06×10^{-16}, 5.5×10^{-16}. All of these are wrong! Let's use series to see what's going on. By Section 13 methods we find, for $x = 0.0015$:

$$\begin{aligned}
\ln\sqrt{(1+x)/(1-x)} &= x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7}\cdots &&= 0.001500001125001518752441,\\
\tan x &= x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315}\cdots &&= 0.001500001125001012500922,\\
f(x) &= \frac{x^5}{15} + \frac{4x^7}{45}\cdots &&= 5.0625 \times 10^{-16}
\end{aligned}$$

with an error of the order of x^7 or 10^{-21}. Now we see that the answer is the difference of two numbers which are identical until the 16th decimal place, so any computer carrying fewer digits will lose all accuracy in the subtraction. It may also be necessary to tell your computer that the value of x is an exact number and not a 4 decimal place approximation. The moral here is that a computer is a tool—a very useful tool, yes—but you need to be constantly aware of whether an answer is reasonable when you are doing problems either by hand or by computer. A final point is that in an applied problem you may want, not a numerical value, but a simple approximation for a complicated function. Here we might approximate $f(x)$ by $x^5/15$ for small x.

FEUE WHEP

Example 2. Evaluate
$$\frac{d^5}{dx^5}\left(\frac{1}{x}\sin x^2\right)\bigg|_{x=0}.$$
We can do this by computer, but it's probably faster to use $\sin x^2 = x^2-(x^2)^3/3!\cdots$, and observe that when we divide this by x and take 5 derivatives, the x^2 term is gone. The second term divided by x is an x^5 term and the fifth derivative of x^5 is 5!. Any further terms will have a power of x which is zero at $x=0$. Thus we have
$$\frac{d^5}{dx^5}\left(\frac{1}{x}\cdot\frac{-(x^2)^3}{3!}\right)_{x=0} = -\frac{5!}{3!} = -20.$$

Summing series We have seen a few numerical series which we could sum exactly (see Sections 1 and 4) and we will see some others later (see Chapter 7, Section 11). Here it is interesting to note that if $f(x)=\sum a_n x^n$, and we let x have a particular value (within the interval of convergence), then we get a numerical series whose sum is the value of the function for that x. For example, if we substitute $x=1$ in (13.4), we get
$$\ln(1+1) = \ln 2 = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}\cdots$$
so the sum of the alternating harmonic series is $\ln 2$.

We can also find sums of series from tables or computer, either the exact sum if that is known, or a numerical approximation (see Problems 20 to 22, and also Problems 14.2, 16.1, 16.30, and 16.31).

Integrals By Theorem 1 of Section 11, we may integrate a power series term by term. Then we can find an approximation for an integral when the indefinite integral cannot be found in terms of elementary functions. As an example, consider the Fresnel integrals (integrals of $\sin x^2$ and $\cos x^2$) which occur in the problem of Fresnel diffraction in optics. We find
$$\begin{aligned}\int_0^t \sin x^2 dx &= \int_0^t \left(x^2-\frac{x^6}{3!}+\frac{x^{10}}{5!}-\cdots\right)dx\\ &= \frac{t^3}{3}-\frac{t^7}{7\cdot 3!}+\frac{t^{11}}{11\cdot 5!}-\cdots\end{aligned}$$
so for $t<1$, the integral is approximately $\frac{t^3}{3}-\frac{t^7}{42}$ with an error <0.00076 since this is an alternating series (see (14.3)).

Evaluation of Indeterminate Forms Suppose we want to find
$$\lim_{x\to 0}\frac{1-e^x}{x}.$$
If we try to substitute $x=0$, we get $0/0$. Expressions that lead us to such meaningless results when we substitute are called indeterminate forms. You can evaluate these by computer, but simple ones can often be done quickly by series. For example,
$$\begin{aligned}\lim_{x\to 0}\frac{1-e^x}{x} &= \lim_{x\to 0}\frac{1-(1+x+(x^2/2!)+\cdots)}{x}\\ &= \lim_{x\to 0}\left(-1-\frac{x}{2!}-\cdots\right) = -1.\end{aligned}$$

FEUE WHEP

You may recall L'Hôpital's rule which says that

$$\lim_{x\to a}\frac{f(x)}{\phi(x)}=\lim_{x\to a}\frac{f'(x)}{\phi'(x)}$$

when $f(a)$ and $\phi(a)$ are both zero, and f'/ϕ' approaches a limit or tends to infinity (that is, does not oscillate) as $x \to a$. Let's use power series to see why this is true. We consider functions $f(x)$ and $\phi(x)$ which are expandable in a Taylor series about $x = a$, and assume that $\phi'(a) \neq 0$. Using (12.8), we have

$$\lim_{x\to a}\frac{f(x)}{\phi(x)}=\lim_{x\to a}\frac{f(a)+(x-a)f'(a)+(x-a)^2 f''(a)/2!+\cdots}{\phi(a)+(x-a)\phi'(a)+(x-a)^2\phi''(a)/2!+\cdots}.$$

If $f(a) = 0$ and $\phi(a) = 0$, and we cancel one $(x - a)$ factor, this becomes

$$\lim_{x\to a}\frac{f'(a)+(x-a)f''(a)/2!+\cdots}{\phi'(a)+(x-a)\phi''(a)/2!+\cdots}=\frac{f'(a)}{\phi'(a)}=\lim_{x\to a}\frac{f'(x)}{\phi'(x)}$$

as L'Hôpital's rule says. If $f'(a) = 0$ and $\phi'(a) = 0$, and $\phi''(a) \neq 0$, then a repetition of the rule gives the limit as $f''(a)/\phi''(a)$, and so on.

There are other indeterminate forms besides $0/0$, for example, ∞/∞, $0\cdot\infty$, etc. L'Hôpital's rule holds for the ∞/∞ form as well as the $0/0$ form. Series are most useful for the $0/0$ form or others which can easily be put into the $0/0$ form. For example, the limit $\lim_{x\to 0}(1/x)\sin x$ is an $\infty\cdot 0$ form, but is easily written as $\lim_{x\to 0}(\sin x)/x$ which is a $0/0$ form. Also *note carefully:* Series (of powers of x) are useful mainly in finding limits as $x \to 0$, because for $x = 0$ such a series collapses to the constant term; for any other value of x we have an infinite series whose sum we probably do not know (see Problem 25, however).

Series Approximations When a problem in, say, differential equations or physics is too difficult in its exact form, we often can get an approximate answer by replacing one or more of the functions in the problem by a few terms of its infinite series. We shall illustrate this idea by two examples.

▸ **Example 3.** In elementary physics we find that the equation of motion of a simple pendulum is (see Chapter 11, Section 8, or a physics textbook):

$$\frac{d^2\theta}{dt^2}=-\frac{g}{l}\sin\theta.$$

This differential equation cannot be solved for θ in terms of elementary functions (see Chapter 11, Section 8), and you may recall that what is usually done is to approximate $\sin\theta$ by θ. Recall the infinite series (13.1) for $\sin\theta$; θ is simply the first term of the series for $\sin\theta$. (Remember that θ is in radians; see discussion in Chapter 2, end of Section 3.) For small values of θ (say $\theta < \frac{1}{2}$ radian or about 30°), this series converges rapidly, and using the first term gives a good approximation (see Problem 14.4). The solutions of the differential equation are then $\theta = A\sin\sqrt{g/l}\,t$ and $\theta = B\cos\sqrt{g/l}\,t$ (A and B constants) as you can verify; we say that the pendulum is executing simple harmonic motion (see Chapter 7, Section 2).

10281

Example 4. Let us consider a radioactive substance containing N_0 atoms at $t = 0$. It is known that the number of atoms remaining at a later time t is given by the formula (see Chapter 8, Section 3):

$$N = N_0 e^{-\lambda t} \tag{15.1}$$

where λ is a constant which is characteristic of the radioactive substance. To find λ for a given substance, a physicist measures in the laboratory the number of decays ΔN during the time interval Δt for a succession of Δt intervals. It is customary to plot each value of $\Delta N/\Delta t$ at the midpoint of the corresponding time interval Δt. If $\lambda\,\Delta t$ is small, this graph is a good approximation to the exact dN/dt graph. A better approximation can be obtained by plotting $\Delta N/\Delta t$ a little to the left of the midpoint. Let us show that the midpoint *does* give a good approximation and also find the more accurate t value. (An approximate value of λ, good enough for calculating the correction, is assumed known from a rough preliminary graph.)

What we should *like* to plot is the graph of dN/dt, that is, the graph of the slope of the curve in Figure 15.1. What we *measure* is the value of $\Delta N/\Delta t$ for each Δt interval. Consider one such Δt interval in Figure 15.1, from t_1 to t_2. To get an accurate graph we should plot the measured value of the quotient $\Delta N/\Delta t$ at the point between t_1 and t_2 where $\Delta N/\Delta t = dN/dt$. Let us write this condition and find the t which satisfies it. The quantity ΔN is the change in N, that is, $N(t_2) - N(t_1)$; the value of dN/dt we get from (15.1). Then $dN/dt = \Delta N/\Delta t$ becomes

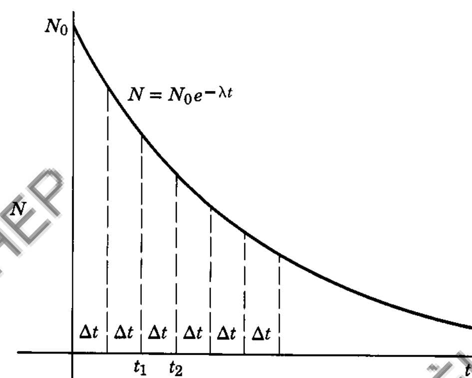


Figure 15.1

$$-\lambda N_0 e^{-\lambda t} = \frac{N_0 e^{-\lambda t_2} - N_0 e^{-\lambda t_1}}{\Delta t}. \tag{15.2}$$

Multiplying this equation by $(\Delta t/N_0)e^{\lambda(t_1+t_2)/2}$, we get

$$-\lambda\,\Delta t\, e^{-\lambda[t-(t_1+t_2)/2]} = e^{-\lambda(t_2-t_1)/2} - e^{\lambda(t_2-t_1)/2} = e^{-\lambda\,\Delta t/2} - e^{\lambda\,\Delta t/2} \tag{15.3}$$

since $t_2 - t_1 = \Delta t$. Since we assumed $\lambda\,\Delta t$ to be small, we can expand the exponentials on the right-hand side of (15.3) in power series; this gives

$$-\lambda\,\Delta t\, e^{-\lambda[t-(t_1+t_2)/2]} = -\lambda\,\Delta t - \frac{1}{3}\left(\frac{\lambda\,\Delta t}{2}\right)^3 \cdots \tag{15.4}$$

or, canceling $(-\lambda\,\Delta t)$,

$$e^{-\lambda[t-(t_1+t_2)/2]} = 1 + \frac{1}{24}(\lambda\,\Delta t)^2 \cdots. \tag{15.5}$$

Suppose $\lambda\,\Delta t$ is small enough so that we can neglect the term $\frac{1}{24}(\lambda\,\Delta t)^2$. Then

10282

(15.5) reduces to

$$e^{-\lambda[t-(t_1+t_2)/2]} = 1,$$
$$-\lambda\left(t - \frac{t_1+t_2}{2}\right) = 0,$$
$$t = \frac{t_1+t_2}{2}.$$

Thus we have justified the usual practice of plotting $\Delta N/\Delta t$ at the midpoint of the interval Δt.

Next consider a more accurate approximation. From (15.5) we get

$$-\lambda\left(t - \frac{t_1+t_2}{2}\right) = \ln\left(1 + \frac{1}{24}(\lambda\,\Delta t)^2 \cdots\right).$$

Since $\frac{1}{24}(\lambda\,\Delta t)^2 \ll 1$, we can expand the logarithm by (13.4) to get

$$-\lambda\left(t - \frac{t_1+t_2}{2}\right) = \frac{1}{24}(\lambda\,\Delta t)^2 \cdots.$$

Then we have

$$t = \frac{t_1+t_2}{2} - \frac{1}{24\lambda}(\lambda\,\Delta t)^2 \cdots.$$

Thus the measured $\Delta N/\Delta t$ should be plotted a little to the left of the midpoint of Δt, as we claimed.

▶ PROBLEMS, SECTION 15

In Problems 1 to 4, use power series to evaluate the function at the given point. Compare with computer results, using the computer to find the series, and also to do the problem without series. Resolve any disagreement in results (see Example 1).

1. $e^{\arcsin x} + \ln\left(\dfrac{1-x}{e}\right)$ at $x = 0.0003$

2. $\dfrac{1}{\sqrt{1+x^4}} - \cos x^2$ at $x = 0.012$

3. $\ln\left(x + \sqrt{1+x^2}\right) - \sin x$ at $x = 0.001$

4. $e^{\sin x} - (1/x^3)\ln(1 + x^3 e^x)$ at $x = 0.00035$

Use Maclaurin series to evaluate each of the following. Although you could do them by computer, you can probably do them in your head faster than you can type them into the computer. So use these to practice quick and skillful use of basic series to make simple calculations.

5. $\dfrac{d^4}{dx^4}\ln(1+x^3)$ at $x = 0$

6. $\dfrac{d^3}{dx^3}\left(\dfrac{x^2 e^x}{1-x}\right)$ at $x = 0$

7. $\dfrac{d^{10}}{dx^{10}}(x^8 \tan^2 x)$ at $x = 0$

10283

8. $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x^2}$ **9.** $\lim\limits_{x\to 0} \dfrac{\sin x - x}{x^3}$ **10.** $\lim\limits_{x\to 0} \dfrac{1-e^{x^3}}{x^3}$

11. $\lim\limits_{x\to 0} \dfrac{\sin^2 2x}{x^2}$ **12.** $\lim\limits_{x\to 0} \dfrac{\tan x - x}{x^3}$ **13.** $\lim\limits_{x\to 0} \dfrac{\ln(1-x)}{x}$

Find a two term approximation for each of the following integrals and an error bound for the given t interval.

14. $\displaystyle\int_0^t e^{-x^2}\,dx, \quad 0 < t < 0.1$ **15.** $\displaystyle\int_0^t \sqrt{x}\,e^{-x}\,dx, \quad 0 < t < 0.01$

Find the sum of each of the following series by recognizing it as the Maclaurin series for a function evaluated at a point.

16. $\displaystyle\sum_{n=1}^{\infty} \frac{2^n}{n!}$ **17.** $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\left(\frac{\pi}{2}\right)^{2n}$

18. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n\,2^n}$ **19.** $\displaystyle\sum_{n=0}^{\infty} \binom{-1/2}{n}\left(-\frac{1}{2}\right)^n$

20. By computer or tables, find the exact sum of each of the following series.

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{n}{(4n^2-1)^2}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{n!}$ (c) $\displaystyle\sum_{n=1}^{\infty} \frac{n(n+1)}{3^n}$

21. By computer, find a numerical approximation for the sum of each of the following series.

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{n}{(n^2+1)^2}$ (b) $\displaystyle\sum_{n=2}^{\infty} \frac{\ln n}{n^2}$ (c) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^n}$

22. The series $\sum_{n=1}^{\infty} 1/n^s$, $s > 1$, is called the Riemann Zeta function, $\zeta(s)$. (In Problem 14.2(a) you found $\zeta(2) = \pi^2/6$. When n is an even integer, these series can be summed exactly in terms of π.) By computer or tables, find

(a) $\zeta(4) = \displaystyle\sum_{n=1}^{\infty} \frac{1}{n^4}$ (b) $\zeta(3) = \displaystyle\sum_{n=1}^{\infty} \frac{1}{n^3}$ (c) $\zeta\left(\dfrac{3}{2}\right) = \displaystyle\sum_{n=1}^{\infty} \frac{1}{n^{3/2}}$

23. Find the following limits using Maclaurin series and check your results by computer. *Hint:* First combine the fractions. Then find the first term of the denominator series and the first term of the numerator series.

(a) $\lim\limits_{x\to 0}\left(\dfrac{1}{x} - \dfrac{1}{e^x-1}\right)$ (b) $\lim\limits_{x\to 0}\left(\dfrac{1}{x^2} - \dfrac{\cos x}{\sin^2 x}\right)$

(c) $\lim\limits_{x\to 0}\left(\csc^2 x - \dfrac{1}{x^2}\right)$ (d) $\lim\limits_{x\to 0}\left(\dfrac{\ln(1+x)}{x^2} - \dfrac{1}{x}\right)$

24. Evaluate the following indeterminate forms by using L'Hôpital's rule and check your results by computer. (Note that Maclaurin series would not be useful here because x does not tend to zero, or because a function ($\ln x$, for example) is not expandable in a Maclaurin series.)

(a) $\lim\limits_{x\to \pi} \dfrac{x\sin x}{x-\pi}$ (b) $\lim\limits_{x\to \pi/2} \dfrac{\ln(2-\sin x)}{\ln(1+\cos x)}$

(c) $\lim\limits_{x\to 1} \dfrac{\ln(2-x)}{x-1}$ (d) $\lim\limits_{x\to \infty} \dfrac{\ln x}{\sqrt{x}}$

(e) $\lim\limits_{x\to 0} x\ln 2x$

FEUE WHEP

10284

(f) $\lim\limits_{x\to\infty} x^n e^{-x}$ (n not necessarily integral)

25. In general, we do not expect Maclaurin series to be useful in evaluating indeterminate forms except when x tends to zero (see Problem 24). Show, however, that Problem 24(f) can be done by writing $x^n e^{-x} = x^n/e^x$ and using the series (13.3) for e^x. *Hint:* Divide numerator and denominator by x^n before you take the limit. What is special about the e^x series which makes it possible to know what the limit of the infinite series is?

26. Find the values of several derivatives of e^{-1/t^2} at $t = 0$. *Hint:* Calculate a few derivatives (as functions of t); then make the substitution $x = 1/t^2$, and use the result of Problem 24(f) or 25.

27. The velocity v of electrons from a high energy accelerator is very near the velocity c of light. Given the voltage V of the accelerator, we often want to calculate the ratio v/c. The relativistic formula for this calculation is (approximately, for $V \gg 1$)

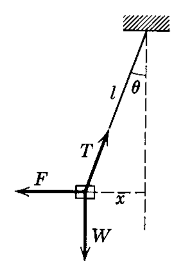
$$\frac{v}{c} = \sqrt{1 - \left(\frac{0.511}{V}\right)^2}, \quad V = \text{number of million volts.}$$

Use two terms of the binomial series (13.5) to find $1 - v/c$ in terms of V. Use your result to find $1 - v/c$ for the following values of V. *Caution:* $V =$ the number of *million* volts.

(a) $V = 100$ million volts
(b) $V = 500$ million volts
(c) $V = 25,000$ million volts
(d) $V = 100$ gigavolts (100×10^9 volts $= 10^5$ million volts)

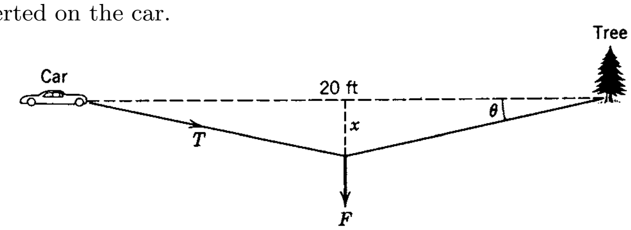
28. The energy of an electron at speed v in special relativity theory is $mc^2(1-v^2/c^2)^{-1/2}$, where m is the electron mass, and c is the speed of light. The factor mc^2 is called the rest mass energy (energy when $v = 0$). Find two terms of the series expansion of $(1 - v^2/c^2)^{-1/2}$, and multiply by mc^2 to get the energy at speed v. What is the second term in the energy series? (If v/c is very small, the rest of the series can be neglected; this is true for everyday speeds.)

29. The figure shows a heavy weight suspended by a cable and pulled to one side by a force F. We want to know how much force F is required to hold the weight in equilibrium at a given distance x to one side (say to place a cornerstone correctly). From elementary physics, $T\cos\theta = W$, and $T\sin\theta = F$.



(a) Find F/W as a series of powers of θ.
(b) Usually in a problem like this, what we know is not θ, but x and l in the diagram. Find F/W as a series of powers of x/l.

30. Given a strong chain and a convenient tree, could you pull your car out of a ditch in the following way? Fasten the chain to the car and to the tree. Pull with a force F at the center of the chain as shown in the figure. From mechanics, we have $F = 2T\sin\theta$, or $T = F/(2\sin\theta)$, where T is the tension in the chain, that is, the force exerted on the car.

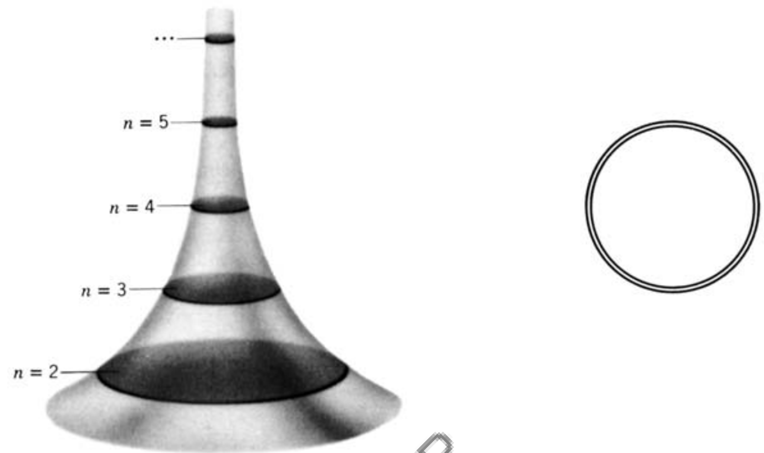


FEUE WHEP
10285

(a) Find T as x^{-1} times a series of powers of x.

(b) Find T as θ^{-1} times a series of powers of θ.

31. A tall tower of circular cross section is reinforced by horizontal circular disks (like large coins), one meter apart and of negligible thickness. The radius of the disk at height n is $1/(n \ln n)$ $(n \geq 2)$.



Assuming that the tower is of infinite height:

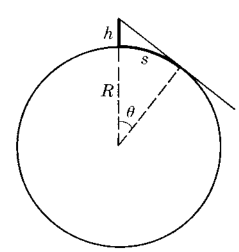
(a) Will the total area of the disks be finite or not? *Hint:* Can you compare the series with a simpler one?

(b) If the disks are strengthened by wires going around their circumferences like tires, will the total length of wire required be finite or not?

(c) Explain why there is not a contradiction between your answers in (a) and (b). That is, how is it possible to start with a set of disks of finite area, remove a little strip around the circumference of each, and get an infinite total length of these strips? *Hint:* Think about units—you can't compare area and length. Consider two cases: (1) Make the width of each strip equal to one percent of the radius of the disk from which you cut it. Now the total length is infinite but what about the total area? (2) Try to make the strips all the same width; what happens? Also see Chapter 5, Problem 3.31(b).

32. Show that the "doubling time" (time for your money to double) is n periods at interest rate $i\%$ per period with $ni = 69$, approximately. Show that the error in the approximation is less than 10% if $i\% \leq 20\%$. (Note that n does not have to be the number of years; it can be the number of months with i = interest rate per month, etc.) *Hint:* You want $(1 + i/100)^n = 2$; take ln of both sides of this equation and use equation (13.4). Also see theorem (14.3).

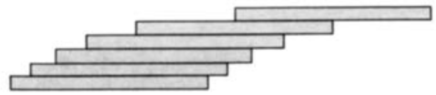
33. If you are at the top of a tower of height h above the surface of the earth, show that the distance you can see along the surface of the earth is approximately $s = \sqrt{2Rh}$, where R is the radius of the earth. *Hints:* See figure. Show that $h/R = \sec\theta - 1$; find two terms of the series for $\sec\theta = 1/\cos\theta$, and use $s = R\theta$. Thus show that the distance in miles is approximately $\sqrt{3h/2}$ with h in feet.



10286

▶ 16. MISCELLANEOUS PROBLEMS

1. (a) Show that it is possible to stack a pile of identical books so that the top book is as far as you like to the right of the bottom book. Start at the top and each time place the pile already completed on top of another book so that the pile is just at the point of tipping. (In practice, of course, you can't let them overhang quite this much without having the stack topple. Try it with

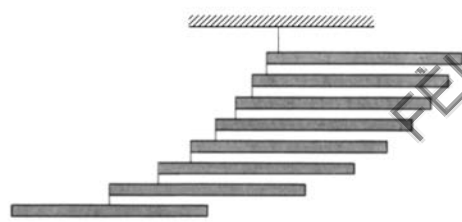


a deck of cards.) Find the distance from the right-hand end of each book to the right-hand end of the one beneath it. To find a general formula for this distance, consider the three forces acting on book n, and write the equation for the torque about its right-hand end. Show that the sum of these setbacks is a divergent series (proportional to the harmonic series). [See "Leaning Tower of *The Physical Reviews,*" Am. J. Phys. **27**, 121–122 (1959).]

(b) By computer, find the sum of N terms of the harmonic series with $N = 25$, 100, 200, 1000, 10^6, 10^{100}.

(c) From the diagram in (a), you can see that with 5 books (count down from the top) the top book is completely to the right of the bottom book, that is, the overhang is slightly over one book. Use your series in (a) to verify this. Then using parts (a) and (b) and a computer as needed, find the number of books needed for an overhang of 2 books, 3 books, 10 books, 100 books.

2. The picture is a mobile constructed of dowels (or soda straws) connected by thin threads. Each thread goes from the left-hand end of a rod to a point on the rod below. Number the rods from the bottom and find, for rod n, the distance from its left end to the thread so that all rods of the mobile will be horizontal. *Hint:* Can you see the relation between this problem and Problem 1?



3. Show that $\sum_{n=2}^{\infty} 1/n^{3/2}$ is convergent. What is wrong with the following "proof" that it diverges?

$$\frac{1}{\sqrt{8}}+\frac{1}{\sqrt{27}}+\frac{1}{\sqrt{64}}+\frac{1}{\sqrt{125}}+\cdots>\frac{1}{\sqrt{9}}+\frac{1}{\sqrt{36}}+\frac{1}{\sqrt{81}}+\frac{1}{\sqrt{144}}+\cdots$$

which is

$$\frac{1}{3}+\frac{1}{6}+\frac{1}{9}+\frac{1}{12}+\cdots=\frac{1}{3}\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots\right).$$

Since the harmonic series diverges, the original series diverges. *Hint:* Compare $3n$ and $n\sqrt{n}$.

Test for convergence:

4. $\displaystyle\sum_{n=1}^{\infty}\frac{2^n}{n!}$ **5.** $\displaystyle\sum_{n=2}^{\infty}\frac{(n-1)^2}{1+n^2}$ **6.** $\displaystyle\sum_{n=2}^{\infty}\frac{\sqrt{n}-1}{(n+1)^2-1}$

7. $\displaystyle\sum_{n=2}^{\infty}\frac{1}{n\ln(n^3)}$ **8.** $\displaystyle\sum_{n=2}^{\infty}\frac{2n^3}{n^4-2}$

10287

Find the interval of convergence, including end-point tests:

9. $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{\ln(n+1)}$ **10.** $\displaystyle\sum_{n=1}^{\infty} \frac{(n!)^2 x^n}{(2n)!}$ **11.** $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n x^{2n-1}}{2n-1}$

12. $\displaystyle\sum_{n=1}^{\infty} \frac{x^n n^2}{5^n(n^2+1)}$ **13.** $\displaystyle\sum_{n=1}^{\infty} \frac{(x+2)^n}{(-3)^n\sqrt{n}}$

Find the Maclaurin series for the following functions.

14. $\cos[\ln(1+x)]$ **15.** $\ln\left(\dfrac{\sin x}{x}\right)$ **16.** $\dfrac{1}{\sqrt{1+\sin x}}$

17. $e^{1-\sqrt{1-x^2}}$ **18.** $\arctan x = \displaystyle\int_0^x \frac{du}{1+u^2}$

Find the first few terms of the Taylor series for the following functions about the given points.

19. $\sin x,\ a=\pi$ **20.** $\sqrt[3]{x},\ a=8$ **21.** $e^x,\ a=1$

Use series you know to show that:

22. $1-\dfrac{1}{3}+\dfrac{1}{5}-\dfrac{1}{7}+\cdots=\dfrac{\pi}{4}$. *Hint:* See Problem 18.

23. $\dfrac{\pi^2}{3!}-\dfrac{\pi^4}{5!}+\dfrac{\pi^6}{7!}-\cdots=1$ **24.** $\ln 3+\dfrac{(\ln 3)^2}{2!}+\dfrac{(\ln 3)^3}{3!}+\cdots=2$

25. Evaluate the limit $\lim_{x\to 0} x^2/\ln\cos x$ by series (in your head), by L'Hôpital's rule, and by computer.

Use Maclaurin series to do Problems 26 to 29 and check your results by computer.

26. $\displaystyle\lim_{x\to 0}\left(\frac{1}{x^2}-\frac{1}{1-\cos^2 x}\right)$ **27.** $\displaystyle\lim_{x\to 0}\left(\frac{1}{x^2}-\cot^2 x\right)$

28. $\displaystyle\lim_{x\to 0}\left(\frac{1+x}{x}-\frac{1}{\sin x}\right)$ **29.** $\displaystyle\left.\frac{d^6}{dx^6}(x^4e^{x^2})\right|_{x=0}$

30. (a) It is clear that you (or your computer) can't find the sum of an infinite series just by adding up the terms one by one. For example, to get $\zeta(1.1)=\sum_{n=1}^{\infty} 1/n^{1.1}$ (see Problem 15.22) with error < 0.005 takes about 10^{33} terms. To see a simple alternative (for a series of positive decreasing terms) look at Figures 6.1 and 6.2. Show that when you have summed N terms, the sum R_N of the rest of the series is between $I_N=\int_N^\infty a_n\,dn$ and $I_{N+1}=\int_{N+1}^\infty a_n\,dn$.

(b) Find the integrals in (a) for the $\zeta(1.1)$ series and verify the claimed number of terms needed for error < 0.005. *Hint*: Find N such that $I_N = 0.005$. Also find upper and lower bounds for $\zeta(1.1)$ by computing $\sum_{n=1}^{N} 1/n^{1.1}+\int_N^\infty n^{-1.1}\,dn$ and $\sum_{n=1}^{N} 1/n^{1.1}+\int_{N+1}^\infty n^{-1.1}\,dn$ where N is far less than 10^{33}. *Hint*: You want the difference between the upper and lower limits to be about 0.005; find N so that term $a_N = 0.005$.

31. As in Problem 30, for each of the following series, find the number of terms required to find the sum with error < 0.005, and find upper and lower bounds for the sum using a much smaller number of terms.

(a) $\displaystyle\sum_{1}^{\infty}\frac{1}{n^{1.01}}$ (b) $\displaystyle\sum_{1}^{\infty}\frac{1}{n(1+\ln n)^2}$ (c) $\displaystyle\sum_{3}^{\infty}\frac{1}{n\ln n(\ln\ln n)^2}$

FEUE WHEP
10288

CHAPTER 2

Complex Numbers

1. INTRODUCTION

You will probably recall using imaginary and complex numbers in algebra. The general solution of the quadratic equation

$$az^2 + bz + c = 0 \tag{1.1}$$

for the unknown z, is given by the *quadratic formula*

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{1.2}$$

If the *discriminant* $d = (b^2 - 4ac)$ is negative, we must take the square root of a negative number in order to find z. Since only non-negative numbers have real square roots, it is impossible to use (1.2) when $d < 0$ unless we introduce a new kind of number, called an imaginary number. We use the symbol $i = \sqrt{-1}$ with the understanding that $i^2 = -1$. Then

$$\sqrt{-16} = 4i, \quad \sqrt{-3} = i\sqrt{3}, \quad i^3 = -i$$

are imaginary numbers, but

$$i^2 = -1, \quad \sqrt{-2}\sqrt{-8} = i\sqrt{2} \cdot i\sqrt{8} = -4, \quad i^{4n} = 1$$

are real. In (1.2) we also need combinations of real and imaginary numbers.

Example. The solution of

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

is

$$z = \frac{2 \pm \sqrt{4-8}}{2} = \frac{2 \pm \sqrt{-4}}{2} = 1 \pm i.$$

We use the term *complex number* to mean any one of the whole set of numbers, real, imaginary, or combinations of the two like $1 \pm i$. Thus, $i + 5$, $17i$, 4, $3 + i\sqrt{5}$ are all examples of complex numbers.

FEUE WHEP
10289

Once the new kind of number is admitted into our number system, fascinating possibilities open up. Can we attach any meaning to marks like $\sin i$, $e^{i\pi}$, $\ln(1+i)$? We'll see later that we can and that, in fact, such expressions may turn up in problems in physics, chemistry, and engineering, as well as mathematics.

When people first considered taking square roots of negative numbers, they felt very uneasy about the problem. They thought that such numbers could not have any meaning or any connection with reality (hence the term "imaginary"). They certainly would not have believed that the new numbers could be of any practical use. Yet complex numbers are of great importance in a variety of applied fields; for example, the electrical engineer would, to say the least, be severely handicapped without them. The complex notation often simplifies setting up and solving vibration problems in either dynamical or electrical systems, and is useful in solving many differential equations which arise from problems in various branches of physics. (See Chapters 7 and 8.) In addition, there is a highly developed field of mathematics dealing with functions of a complex variable (see Chapter 14) which yields many useful methods for solving problems about fluid flow, elasticity, quantum mechanics, and other applied problems. Almost every field of either pure or applied mathematics makes some use of complex numbers.

2. REAL AND IMAGINARY PARTS OF A COMPLEX NUMBER

A complex number such as $5 + 3i$ is the sum of two terms. The real term (not containing i) is called the *real part* of the complex number. The *coefficient* of i in the other term is called the *imaginary part* of the complex number. In $5 + 3i$, 5 is the real part and 3 is the imaginary part. Notice carefully that the *imaginary part* of a complex number is *not imaginary*!

Either the real part or the imaginary part of a complex number may be zero. If the real part is zero, the complex number is called imaginary (or, for emphasis, *pure* imaginary). The zero real part is usually omitted; thus $0 + 5i$ is written just $5i$. If the imaginary part of the complex number is zero, the number is real. We write $7 + 0i$ as just 7. Complex numbers then include both real numbers and pure imaginary numbers as special cases.

In algebra a complex number is ordinarily written (as we have been doing) as a sum like $5+3i$. There is another very useful way of thinking of a complex number. As we have said, every complex number has a real part and an imaginary part (either of which may be zero). These are two *real* numbers, and we could, if we liked, agree to write $5+3i$ as $(5, 3)$. Any complex number could be written this way as a pair of real numbers, the real part first and then the imaginary part (which, you must remember, is real). This would not be a very convenient form for computation, but it suggests a very useful geometrical representation of a complex number which we shall now consider.

3. THE COMPLEX PLANE

In analytic geometry we plot the point $(5, 3)$ as shown in Figure 3.1. As we have seen, the symbol $(5, 3)$ could also mean the complex number $5+3i$. The point $(5, 3)$ may then be labeled either $(5, 3)$ or $5+3i$. Similarly, any complex number $x+iy$ (x and y real) can be represented by a point (x, y) in the (x, y) plane. Also any point (x, y) in the (x, y) plane can be labeled $x + iy$ as well as (x, y). When the (x, y)

10290

plane is used in this way to plot complex numbers, it is called the *complex plane*. It is also sometimes called an *Argand diagram*. The x axis is called the real axis, and the y axis is called the imaginary axis (note, however, that you plot y and *not* iy).

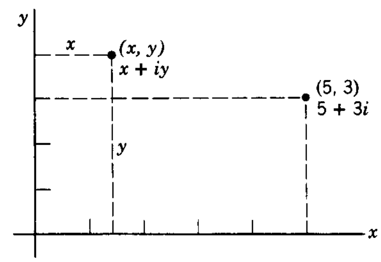


Figure 3.1

When a complex number is written in the form $x + iy$, we say that it is in *rectangular form* because x and y are the rectangular coordinates of the point representing the number in the complex plane. In analytic geometry, we can locate a point by giving its polar coordinates (r, θ) instead of its rectangular coordinates (x, y). There is a corresponding way to write any complex number. In Figure 3.2,

$$\begin{aligned} x &= r\cos\theta, \\ y &= r\sin\theta. \end{aligned} \tag{3.1}$$

Then we have

$$\begin{aligned} x + iy &= r\,\cos\theta + ir\,\sin\theta \\ &= r\,(\cos\theta + i\,\sin\theta). \end{aligned} \tag{3.2}$$

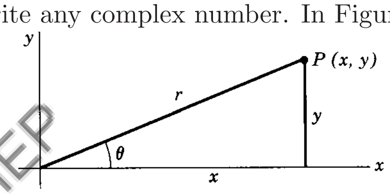


Figure 3.2

This last expression is called the *polar form* of the complex number. As we shall see (Sections 9 to 16), the expression $(\cos\theta + i\,\sin\theta)$ can be written as $e^{i\theta}$, so a convenient way to write the polar form of a complex number is

$$x + iy = r(\cos\theta + i\,\sin\theta) = re^{i\theta}. \tag{3.3}$$

The polar form $re^{i\theta}$ of a complex number is often simpler to use than the rectangular form.

▶ **Example.** In Figure 3.3 the point A could be labeled as $(1, \sqrt{3})$ or as $1 + i\sqrt{3}$. Similarly, using polar coordinates, the point A could be labeled with its (r, θ) values as $(2, \pi/3)$. Notice that r is always taken positive. Using (3.3) we have

$$1 + i\sqrt{3} = 2\left(\cos\frac{\pi}{3} + i\,\sin\frac{\pi}{3}\right) = 2e^{i\pi/3}.$$

This gives two more ways to label point A in Figure 3.3.

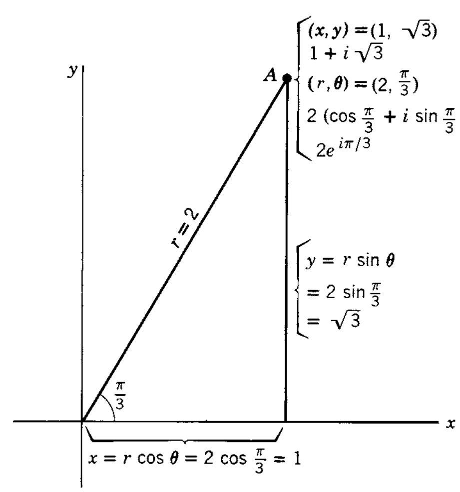


Figure 3.3

FÊUE WHEP
10291

Radians and Degrees In Figure 3.3, the angle $\pi/3$ is in radians. Ever since you studied calculus, you have been expected to measure angles in radians and not degrees. Do you know why? You have learned that $(d/dx)\sin x = \cos x$. This formula is **not** correct—**unless** x is in radians. (Look up the derivation in your calculus book!) Many of the formulas you now know and use are correct *only* if you use radian measure; consequently that is what you are usually advised to do. However, it is sometimes convenient to do computations with complex numbers using degrees, so it is important to know when you can and when you cannot use degrees. You can use degrees to measure an angle and to add and subtract angles as long as the final step is to find the sine, cosine, or tangent of the resulting angle (with your calculator in degree mode). For example, in Figure 3.3, we can, if we like, say that $\theta = 60°$ instead of $\theta = \pi/3$. If we want to find $\sin(\pi/3 - \pi/4) = \sin(\pi/12) = 0.2588$ (calculator in radian mode), we can instead find $\sin(60° - 45°) = \sin 15° = 0.2588$ (calculator in degree mode). Note carefully that an angle is in radians unless the degree symbol is used; for example, in $\sin 2$, the 2 is 2 radians or about 115°.

In formulas, however, use radians. For example, in using infinite series, we say that $\sin\theta \cong \theta$ for very small θ. Try this on your calculator; you will find that it is true in radian mode but not in degree mode. As another example, consider $\int_0^1 dx/(1+x^2) = \arctan 1 = \pi/4 = 0.785$. Here $\arctan 1$ is *not* an angle; it is the numerical value of the integral, so the answer 45 (obtained from a calculator in degree mode) is wrong! Do not use degree mode in reading an arc tan (or arc sin or arc cos) *unless* you are finding an *angle* [for example, in Figure 3.2, $\theta = \arctan(y/x)$, and in Figure 3.3, $\theta = \arctan\sqrt{3} = \pi/3$ or 60°].

4. TERMINOLOGY AND NOTATION

Both i and j are used to represent $\sqrt{-1}$, j usually in any problem dealing with electricity since i is needed there for current. A physicist should be able to work with ease using either symbol. We shall for consistency use i throughout this book.

We often label a point with a single letter (for example, P in Figure 3.2 and A in Figure 3.3) even though it requires two coordinates to locate the point. If you have studied vectors, you will recall that a vector is represented by a single letter, say $\mathbf{v}$, although it has (in two dimensions) two components. It is customary to use a single letter for a complex number even though we realize that it is actually a pair of real numbers. Thus we write

$$z = x + iy = r(\cos\theta + i\sin\theta) = re^{i\theta}. \tag{4.1}$$

Here z is a complex number; x is the *real part* of the complex number z, and y is the *imaginary part* of z. The quantity r is called the *modulus* or *absolute value* of z, and θ is called the *angle* of z (or the *phase*, or the *argument*, or the *amplitude* of z). In symbols:

$$\begin{aligned} &\operatorname{Re} z = x, && |z| = \operatorname{mod} z = r = \sqrt{x^2+y^2}, \\ &\operatorname{Im} z = y\ (\textbf{not}\ iy), && \text{angle of } z = \theta. \end{aligned} \tag{4.2}$$

The values of θ should be found from a diagram rather than a formula, although we do sometimes write $\theta = \arctan(y/x)$. An example shows this clearly.

▶ **Example.** Write $z = -1 - i$ in polar form. Here we have $x = -1$, $y = -1$, $r = \sqrt{2}$ (Figure 4.1). There are an infinite number of values of θ,

$$\theta = \frac{5\pi}{4} + 2n\pi, \tag{4.3}$$

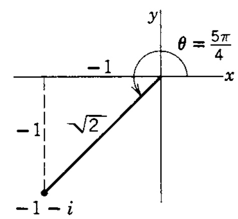


Figure 4.1

where n is any integer, positive or negative. The value $\theta = 5\pi/4$ is sometimes called the *principal angle* of the complex number $z = -1 - i$. Notice carefully, however, that this is not the same as the principal value $\pi/4$ of $\arctan 1$ as defined in calculus. The angle of a complex number must be in the same quadrant as the point representing the number. For our present work, any one of the values in (4.3) will do; here we would probably use either $5\pi/4$ or $-3\pi/4$. Then we have in our example

$$\begin{aligned} z = -1 - i &= \sqrt{2}\left[\cos\left(\frac{5\pi}{4} + 2n\pi\right) + i\sin\left(\frac{5\pi}{4} + 2n\pi\right)\right] \\ &= \sqrt{2}\left(\cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4}\right) = \sqrt{2}\, e^{5i\pi/4}. \end{aligned}$$

[We could also write $z = \sqrt{2}\,(\cos 225° + i \sin 225°)$.]

The complex number $x - iy$, obtained by changing the sign of i in $z = x + iy$, is called the *complex conjugate* or simply the *conjugate* of z. We usually write the conjugate of $z = x + iy$ as $\bar{z} = x - iy$. Sometimes we use z^* instead of $\bar{z}$ (in fields such as statistics or quantum mechanics where the bar may be used to mean an average value). Notice carefully that the conjugate of $7i - 5$ is $-7i - 5$; that is, it is the i term whose sign is changed.

Complex numbers come in conjugate pairs; for example, the conjugate of $2 + 3i$ is $2 - 3i$ and the conjugate of $2 - 3i$ is $2 + 3i$. Such a pair of points in the complex plane are mirror images of each other with the x axis as the mirror (Figure 4.2). Then in polar form, z and $\bar{z}$ have the same r value, but their θ values are negatives of each other. If we write $z = r(\cos\theta + i\sin\theta)$, then

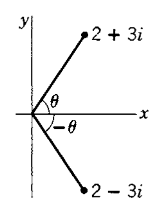


Figure 4.2

$$\bar{z} = r[\cos(-\theta) + i\sin(-\theta)] = r(\cos\theta - i\sin\theta) = re^{-i\theta}. \tag{4.4}$$

▶ PROBLEMS, SECTION 4

For each of the following numbers, first visualize where it is in the complex plane. With a little practice you can quickly find x, y, r, θ in your head for these simple problems. Then

10293

plot the number and label it in five ways as in Figure 3.3. Also plot the complex conjugate of the number.

1. $1+i$ **2.** $i-1$ **3.** $1-i\sqrt{3}$

4. $-\sqrt{3}+i$ **5.** $2i$ **6.** $-4i$

7. -1 **8.** 3 **9.** $2i-2$

10. $2-2i$ **11.** $2\left(\cos\frac{\pi}{6}+i\sin\frac{\pi}{6}\right)$

12. $4\left(\cos\frac{2\pi}{3}-i\sin\frac{2\pi}{3}\right)$ **13.** $\cos\frac{3\pi}{2}+i\sin\frac{3\pi}{2}$

14. $2\left(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4}\right)$ **15.** $\cos\pi-i\sin\pi$

16. $5(\cos 0+i\sin 0)$ **17.** $\sqrt{2}\,e^{-i\pi/4}$

18. $3\,e^{i\pi/2}$ **19.** $5(\cos 20^\circ+i\sin 20^\circ)$

20. $7(\cos 110^\circ-i\sin 110^0)$

5. COMPLEX ALGEBRA

A. Simplifying to $x+iy$ form

Any complex number can be written in the rectangular form $x+iy$. To add, subtract, and multiply complex numbers, remember that they follow the ordinary rules of algebra and that $i^2=-1$.

Example 1.

$$(1+i)^2=1+2i+i^2=1+2i-1=2i$$

To divide one complex number by another, first write the quotient as a fraction. Then reduce the fraction to rectangular form by multiplying numerator and denominator by the conjugate of the denominator; this makes the denominator real.

Example 2.

$$\frac{2+i}{3-i}=\frac{2+i}{3-i}\cdot\frac{3+i}{3+i}=\frac{6+5i+i^2}{9-i^2}=\frac{5+5i}{10}=\frac{1}{2}+\frac{1}{2}i.$$

It is sometimes easier to multiply or divide complex numbers in polar form.

Example 3. To find $(1+i)^2$ in polar form, we first sketch (or picture mentally) the point $(1,1)$. From Figure 5.1, we see that $r=\sqrt{2}$, and $\theta=\pi/4$, so $(1+i)=\sqrt{2}\,e^{i\pi/4}$. Then from Figure 5.2 we find the same result as in Example 1.

$$(1+i)^2=(\sqrt{2}\,e^{i\pi/4})^2=2\,e^{i\pi/2}=2i.$$

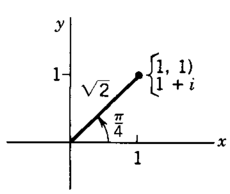


Figure 5.1

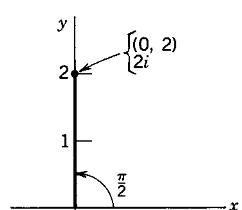


Figure 5.2

▶ **Example 4.** Write $1/[2(\cos 20° + i\sin 20°]$ in $x+iy$ form. Since $20° = \pi/9$ radians,

$$\frac{1}{2(\cos 20° + i\sin 20°)} = \frac{1}{2(\cos\pi/9 + i\sin\pi/9)} = \frac{1}{2\,e^{i\pi/9}} = 0.5\,e^{-i\pi/9}$$
$$= 0.5(\cos\pi/9 - i\sin\pi/9) = 0.47 - 0.17i,$$

by calculator in radian mode. We obtain the same result leaving the angle in degrees and using a calculator in degree mode: $0.5(\cos 20° - i\sin 20°) = 0.47 - 0.17i$.

▶ PROBLEMS, SECTION 5

First simplify each of the following numbers to the $x+iy$ form or to the $re^{i\theta}$ form. Then plot the number in the complex plane.

1. $\dfrac{1}{1+i}$ **2.** $\dfrac{1}{i-1}$ **3.** i^4

4. i^2+2i+1 **5.** $\left(i+\sqrt{3}\right)^2$ **6.** $\left(\dfrac{1+i}{1-i}\right)^2$

7. $\dfrac{3+i}{2+i}$ **8.** $1.6-2.7i$

9. $25\,e^{2i}$ *Careful!* The angle is 2 radians.

10. $\dfrac{3i-7}{i+4}$ *Careful! Not* $3-7i$

11. $17-12i$ **12.** $3(\cos 28° + i\sin 28°)$

13. $5\left(\cos\dfrac{2\pi}{5} + i\sin\dfrac{2\pi}{5}\right)$ **14.** $2.8e^{-i(1.1)}$

15. $\dfrac{5-2i}{5+2i}$ **16.** $\dfrac{1}{0.5(\cos 40° + i\sin 40°)}$

17. $(1.7-3.2i)^2$ **18.** $(0.64+0.77i)^4$

Find each of the following in rectangular $(a+bi)$ form if $z = 2-3i$; if $z = x+iy$.

19. z^{-1} **20.** $\dfrac{1}{z^2}$ **21.** $\dfrac{1}{z+1}$

22. $\dfrac{1}{z-i}$ **23.** $\dfrac{1+z}{1-z}$ **24.** $z/\bar{z}$

B. Complex Conjugate of a Complex Expression

It is easy to see that the conjugate of the sum of two complex numbers is the sum of the conjugates of the numbers. If

$$z_1 = x_1 + iy_1 \qquad \text{and} \qquad z_2 = x_2 + iy_2,$$

then

$$\bar{z}_1 + \bar{z}_2 = x_1 - iy_1 + x_2 - iy_2 = x_1 + x_2 - i(y_1+y_2).$$

The conjugate of (z_1+z_2) is

$$\overline{(x_1+x_2) + i(y_1+y_2)} = (x_1+x_2) - i(y_1+y_2).$$

FÊUE WHEP

Similarly, you can show that the conjugate of the difference (or product or quotient) of two complex numbers is equal to the difference (or product or quotient) of the conjugates of the numbers (Problem 25). In other words, you can get the conjugate of an expression containing i's by just changing the signs of all the i terms. We must watch out for hidden i's, however.

Example. If

$$z = \frac{2-3i}{i+4}, \qquad \text{then} \qquad \bar{z} = \frac{2+3i}{-i+4}.$$

But if $z = f + ig$, where f and g are themselves complex, then the complex conjugate of z is $\bar{z} = \bar{f} - i\bar{g}$ (*not* $f - ig$).

PROBLEMS, SECTION 5

25. Prove that the conjugate of the quotient of two complex numbers is the quotient of the conjugates. Also prove the corresponding statements for difference and product. *Hint:* It is easier to prove the statements about product and quotient using the polar coordinate $re^{i\theta}$ form; for the difference, it is easier to use the rectangular form $x + iy$.

C. Finding the Absolute Value of z

Recall that the definition of $|z|$ is $|z| = r = \sqrt{x^2+y^2}$ (positive square root!). Since $z\bar{z} = (x+iy)(x-iy) = x^2+y^2$, or, in polar coordinates, $z\bar{z} = (re^{i\theta})(re^{-i\theta}) = r^2$, we see that $|z|^2 = z\bar{z}$, or $|z| = \sqrt{z\bar{z}}$. Note that $z\bar{z}$ is always real and ≥ 0, since x, y, and r are real. We have

$$|z| = r = \sqrt{x^2+y^2} = \sqrt{z\bar{z}}. \tag{5.1}$$

By Problem 25 and (5.1), the absolute value of a quotient of two complex numbers is the quotient of the absolute values (and a similar statement for product).

Example.

$$\left|\frac{\sqrt{5}+3i}{1-i}\right| = \frac{|\sqrt{5}+3i|}{|1-i|} = \frac{\sqrt{14}}{\sqrt{2}} = \sqrt{7}.$$

PROBLEMS, SECTION 5

Find the absolute value of each of the following using the discussion above. Try to do simple problems like these in your head—it saves time.

26. $\dfrac{2i-1}{i-2}$ **27.** $\dfrac{2+3i}{1-i}$ **28.** $\dfrac{z}{\bar{z}}$

29. $(1+2i)^3$ **30.** $\dfrac{3i}{i-\sqrt{3}}$ **31.** $\dfrac{5-2i}{5+2i}$

32. $(2-3i)^4$ **33.** $\dfrac{25}{3+4i}$ **34.** $\left(\dfrac{1+i}{1-i}\right)^5$

FEUE WHEP

D. Complex Equations

In working with equations involving complex quantities, we must always remember that a complex number is actually a pair of real numbers. Two complex numbers are equal if and only if their real parts are equal and their imaginary parts are equal. For example, $x + iy = 2 + 3i$ means $x = 2$ and $y = 3$. In other words, any equation involving complex numbers is really two equations involving real numbers.

▶ **Example.** Find x and y if

$$(x + iy)^2 = 2i. \tag{5.2}$$

Since $(x + iy)^2 = x^2 + 2ixy - y^2$, (5.2) is equivalent to the two real equations

$$x^2 - y^2 = 0,$$
$$2xy = 2.$$

From the first equation $y^2 = x^2$, we find $y = x$ or $y = -x$. Substituting these into the second equation gives

$$2x^2 = 2 \qquad \text{or} \qquad -2x^2 = 2.$$

Since x is real, x^2 cannot be negative. Thus we find only

$$x^2 = 1 \qquad \text{and} \qquad y = x,$$

that is,

$$x = y = 1 \qquad \text{and} \qquad x = y = -1.$$

▶ PROBLEMS, SECTION 5

Solve for all possible values of the real numbers x and y in the following equations.

35. $x + iy = 3i - 4$

36. $2ix + 3 = y - i$

37. $x + iy = 0$

38. $x + iy = 2i - 7$

39. $x + iy = y + ix$

40. $x + iy = 3i - ix$

41. $(2x - 3y - 5) + i(x + 2y + 1) = 0$

42. $(x + 2y + 3) + i(3x - y - 1) = 0$

43. $(x + iy)^2 = 2ix$

44. $x + iy = (1 - i)^2$

45. $(x + iy)^2 = (x - iy)^2$

46. $\dfrac{x + iy}{x - iy} = -i$

47. $(x + iy)^3 = -1$

48. $\dfrac{x + iy + 2 + 3i}{2x + 2iy - 3} = i + 2$

49. $|1 - (x + iy)| = x + iy$

50. $|x + iy| = y - ix$

E. Graphs

Using the graphical representation of the complex number z as the point (x, y) in a plane, we can give geometrical meaning to equations and inequalities involving z.

FEUE WHEP

Example 1. What is the curve made up of the points in the (x, y) plane satisfying the equation $|z| = 3$?

Since

$$|z| = \sqrt{x^2 + y^2},$$

the given equation is

$$\sqrt{x^2 + y^2} = 3 \quad \text{or} \quad x^2 + y^2 = 9.$$

Thus $|z| = 3$ is the equation of a circle of radius 3 with center at the origin. Such an equation might describe, for example, the path of an electron or of a satellite. (See Section F below.)

Example 2.

(a) $|z - 1| = 2$. This is the circle $(x - 1)^2 + y^2 = 4$.
(b) $|z - 1| \le 2$. This is the disk whose boundary is the circle in (a).

Note that we use "circle" to mean a curve and "disk" to mean an area. The interior of the disk is given by $|z - 1| < 2$.

Example 3. (Angle of z) $= \pi/4$. This is the half-line $y = x$ with $x > 0$; this might be the path of a light ray starting at the origin.

Example 4. $\operatorname{Re} z > \frac{1}{2}$. This is the half-plane $x > \frac{1}{2}$.

PROBLEMS, SECTION 5

Describe geometrically the set of points in the complex plane satisfying the following equations.

51. $|z| = 2$

52. $\operatorname{Re} z = 0$

53. $|z - 1| = 1$

54. $|z - 1| < 1$

55. $z - \bar{z} = 5i$

56. angle of $z = \dfrac{\pi}{2}$

57. $\operatorname{Re}(z^2) = 4$

58. $\operatorname{Re} z > 2$

59. $|z + 3i| = 4$

60. $|z - 1 + i| = 2$

61. $\operatorname{Im} z < 0$

62. $|z + 1| + |z - 1| = 8$

63. $z^2 = \bar{z}^2$

64. $z^2 = -\bar{z}^2$

65. Show that $|z_1 - z_2|$ is the distance between the points z_1 and z_2 in the complex plane. Use this result to identify the graphs in Problems 53, 54, 59, and 60 without computation.

F. Physical Applications

Problems in physics as well as geometry may often be simplified by using one complex equation instead of two real equations. See the following example and also Section 16.

FEUE WHEP

▶ **Example.** A particle moves in the (x, y) plane so that its position (x, y) as a function of time t is given by

$$z = x + iy = \frac{i + 2t}{t - i}.$$

Find the magnitudes of its velocity and its acceleration as functions of t.

We *could* write z in $x + iy$ form and so find x and y as functions of t. It is easier to do the problem as follows. We define the complex velocity and complex acceleration by

$$\frac{dz}{dt} = \frac{dx}{dt} + i\,\frac{dy}{dt} \quad \text{and} \quad \frac{d^2z}{dt^2} = \frac{d^2x}{dt^2} + i\,\frac{d^2y}{dt^2}.$$

Then the magnitude v of the velocity is $v = \sqrt{(dx/dt)^2 + (dy/dt)^2} = |dz/dt|$, and similarly the magnitude a of the acceleration is $a = |d^2z/dt^2|$. Thus we have

$$\frac{dz}{dt} = \frac{2(t - i) - (i + 2t)}{(t - i)^2} = \frac{-3i}{(t - i)^2}.$$

$$v = \left|\frac{dz}{dt}\right| = \sqrt{\frac{-3i}{(t - i)^2} \cdot \frac{+3i}{(t + i)^2}} = \frac{3}{t^2 + 1},$$

$$\frac{d^2z}{dt^2} = \frac{(-3i)(-2)}{(t - i)^3} = \frac{6i}{(t - i)^3},$$

$$a = \left|\frac{d^2z}{dt^2}\right| = \frac{6}{(t^2 + 1)^{3/2}}.$$

Note carefully that all physical quantities (x, y, v, and a) are real; the complex expressions are used just for convenience in calculation.

▶ PROBLEMS, SECTION 5

66. Find x and y as functions of t for the example above, and verify for this case that v and a are correctly given by the method of the example.

67. Find v and a if $z = (1 - it)/(2t + i)$.

68. Find v and a if $z = \cos 2t + i \sin 2t$. Can you describe the motion?

▶ 6. COMPLEX INFINITE SERIES

In Chapter 1 we considered infinite series whose terms were real. We shall be very much interested in series with complex terms; let us reconsider our definitions and theorems for this case. The partial sums of a series of complex numbers will be complex numbers, say $S_n = X_n + iY_n$, where X_n and Y_n are real. Convergence is defined just as for real series: If S_n approaches a limit $S = X + iY$ as $n \to \infty$, we call the series convergent and call S its sum. This means that $X_n \to X$ and $Y_n \to Y$; in other words, the real and the imaginary parts of the series are each convergent series.

It is useful, just as for real series, to discuss absolute convergence first. It can be proved (Problem 1) that an absolutely convergent series converges. Absolute convergence means here, just as for real series, that the series of absolute values of the terms is a convergent series. Remember that $|z| = r = \sqrt{x^2 + y^2}$ is a positive number. Thus any of the tests given in Chapter 1 for convergence of series of positive terms may be used here to test a complex series for absolute convergence.

FEUE WHEP

Example 1. Test for convergence

$$1 + \frac{1+i}{2} + \frac{(1+i)^2}{4} + \frac{(1+i)^3}{8} + \cdots + \frac{(1+i)^n}{2^n} + \cdots.$$

Using the ratio test, we find

$$\rho = \lim_{n\to\infty}\left|\frac{(1+i)^{n+1}}{2^{n+1}} \div \frac{(1+i)^n}{2^n}\right| = \lim_{n\to\infty}\left|\frac{1+i}{2}\right| = \left|\frac{1+i}{2}\right| = \frac{\sqrt{2}}{2} < 1.$$

. Since $\rho < 1$, the series is absolutely convergent and therefore convergent.

Example 2. Test for convergence $\sum_1^\infty i^n/\sqrt{n}$. Here the ratio test gives 1 so we must try a different test. Let's write out a few terms of the series:

$$i - \frac{1}{\sqrt{2}} - \frac{i}{\sqrt{3}} + \frac{1}{\sqrt{4}} + \frac{i}{\sqrt{5}} - \frac{1}{\sqrt{6}} \cdots.$$

We see that the real part of the series is

$$-\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{4}} - \frac{1}{\sqrt{6}} + \cdots = \sum_1^\infty \frac{(-1)^n}{\sqrt{2n}},$$

and the imaginary part of the series is

$$1 - \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{5}} \cdots = \sum_0^\infty \frac{(-1)^n}{\sqrt{2n+1}}.$$

Verify that both these series satisfy the alternating series test for convergence. Thus, the original series converges.

Example 3. Test for convergence $\sum_0^\infty z^n = \sum_0^\infty (re^{i\theta})^n = \sum_0^\infty r^n e^{in\theta}$. This is a geometric series with ratio $= z = re^{i\theta}$; it converges if and only if $|z| < 1$. Recall that $|z| = r$. Thus, $\sum_0^\infty r^n e^{in\theta}$ converges if and only if $r < 1$.

PROBLEMS, SECTION 6

1. Prove that an absolutely convergent series of complex numbers converges. This means to prove that $\sum(a_n + ib_n)$ converges (a_n and b_n real) if $\sum\sqrt{a_n^2 + b_n^2}$ converges. *Hint:* Convergence of $\sum(a_n+ib_n)$ means that $\sum a_n$ and $\sum b_n$ *both* converge. Compare $\sum|a_n|$ and $\sum|b_n|$ with $\sum\sqrt{a_n^2+b_n^2}$, and use Problem 7.9 of Chapter 1.

Test each of the following series for convergence.

2. $\sum(1+i)^n$ **3.** $\sum\frac{1}{(1+i)^n}$ **4.** $\sum\left(\frac{1-i}{1+i}\right)^n$

5. $\sum\left(\frac{1}{n^2}+\frac{i}{n}\right)$ **6.** $\sum\frac{1+i}{n^2}$ **7.** $\sum\frac{(i-1)^n}{n}$

8. $\sum e^{in\pi/6}$ **9.** $\sum\frac{i^n}{n}$ **10.** $\sum\left(\frac{1+i}{1-i\sqrt{3}}\right)^n$

11. $\sum\left(\frac{2+i}{3-4i}\right)^{2n}$ **12.** $\sum\frac{(3+2i)^n}{n!}$ **13.** $\sum\left(\frac{1+i}{2-i}\right)^n$

14. Prove that a series of complex terms diverges if $\rho > 1$ (ρ = ratio test limit). *Hint:* The nth term of a convergent series tends to zero.

FÉUE WHEP
10300

▶ 7. COMPLEX POWER SERIES; DISK OF CONVERGENCE

In Chapter 1 we considered series of powers of x, $\sum a_n x^n$. We are now interested in series of powers of z,

$$\sum a_n z^n, \tag{7.1}$$

where $z = x + iy$, and the a_n are complex numbers. [Notice that (7.1) includes real series as a special case since $z = x$ if $y = 0$.] Here are some examples.

$$1 - z + \frac{z^2}{2} - \frac{z^3}{3} + \frac{z^4}{4} + \cdots, \tag{7.2a}$$

$$1 + iz + \frac{(iz)^2}{2!} + \frac{(iz)^3}{3!} + \cdots = 1 + iz - \frac{z^2}{2!} - \frac{iz^3}{3!} + \cdots, \tag{7.2b}$$

$$\sum_{n=0}^{\infty} \frac{(z+1-i)^n}{3^n\, n^2}. \tag{7.2c}$$

Let us use the ratio test to find for what z these series are absolutely convergent. For (7.2a), we have

$$\rho = \lim_{n\to\infty} \left| \frac{z \cdot n}{n+1} \right| = |z|.$$

The series converges if $\rho < 1$, that is, if $|z| < 1$, or $\sqrt{x^2+y^2} < 1$. This is the interior of a disk of radius 1 with center at the origin in the complex plane. This disk is called the *disk of convergence* of the infinite series and the radius of the disk is called the *radius of convergence*. The disk of convergence replaces the interval of convergence which we had for real series. In fact (see Figure 7.1), the interval of convergence for the series $\sum (-x)^n/n$ is just the interval $(-1, 1)$ on the x axis contained within the disk of convergence of $\sum (-z)^n/n$, as it must be since x is the value of z when $y = 0$. For this reason we sometimes speak of the *radius* of convergence of a power series even though we are considering only real values of z. (Also see Chapter 14, Equations (2.5) and (2.6) and Figure 2.4.)

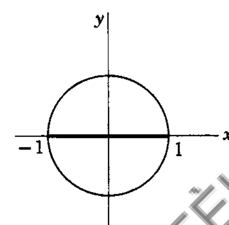


Figure 7.1

Next consider series (7.2b); here we have

$$\rho = \lim_{n\to\infty} \left| \frac{(iz)^{n+1}}{(n+1)!} \div \frac{(iz)^n}{n!} \right| = \lim_{n\to\infty} \left| \frac{iz}{n+1} \right| = 0.$$

This is an example of a series which converges for all values of z. For series (7.2c), we have

$$\rho = \lim_{n\to\infty} \left| \frac{(z+1-i)}{3} \frac{n^2}{(n+1)^2} \right| = \left| \frac{z+1-i}{3} \right|.$$

Thus, this series converges for

$$|z+1-i| < 3, \text{ or } |z-(-1+i)| < 3.$$

This is the interior of a disk (Figure 7.2) of radius 3 and center at $z = -1 + i$ (see Problem 5.65).

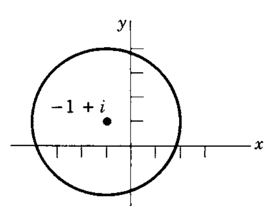


Figure 7.2

FEUE WHEP
10301

Just as for real series, if $\rho > 1$, the series diverges (Problem 6.14). For $\rho = 1$ (that is, on the boundary of the disk of convergence) the series may either converge or diverge. It may be difficult to find out which and we shall not in general need to consider the question.

The four theorems about power series (Chapter 1, Section 11) are true also for complex series (replace *interval* by *disk* of convergence). Also we can now state for Theorem 2 what the disk of convergence is for the quotient of two series of powers of z. Assume to start with that any common factor z has been cancelled. Let r_1 and r_2 be the radii of convergence of the numerator and denominator series. Find the closest point to the origin in the complex plane where the denominator is zero; call the distance from the origin to this point s. Then the quotient series converges at least inside the smallest of the three disks of radii r_1, r_2, and s, with center at the origin. (See Chapter 14, Section 2.)

▶ **Example.** Find the disk of convergence of the Maclaurin series for $(\sin z)/[z(1+z^2)]$.

We shall soon see that the series for $\sin z$ has the same form as the real series for $\sin x$ in Chapter 1. Using this fact we find (Problem 17)

$$\frac{\sin z}{z(1+z^2)} = 1 - \frac{7z^2}{6} + \frac{47z^4}{40} - \frac{5923z^6}{5040} + \cdots . \tag{7.3}$$

From (7.3) we can't find the radius of convergence, but let's use the theorem above. Let the numerator series be $(\sin z)/z$. By ratio test, the series for $(\sin z)/z$ converges for all z (if you like, $r_1 = \infty$). There is no r_2 since the denominator is not an infinite series. The denominator $1 + z^2$ is zero when $z = \pm i$, so $s = 1$. Then the series (7.3) converges inside a disk of radius 1 with center at the origin.

▶ PROBLEMS, SECTION 7

Find the disk of convergence for each of the following complex power series.

1. $e^z = 1 + z + \dfrac{z^2}{2!} + \dfrac{z^3}{3!} \cdots$ [equation (8.1)]

2. $z - \dfrac{z^2}{2} + \dfrac{z^3}{3} - \dfrac{z^4}{4} + \cdots$

3. $1 - \dfrac{z^2}{3!} + \dfrac{z^4}{5!} - \cdots$

4. $\displaystyle\sum_{n=0}^{\infty} z^n$

5. $\displaystyle\sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^n$

6. $\displaystyle\sum_{n=1}^{\infty} n^2 (3iz)^n$

7. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!}$

8. $\displaystyle\sum_{n=1}^{\infty} \frac{z^{2n}}{(2n+1)!}$

9. $\displaystyle\sum_{n=1}^{\infty} \frac{z^n}{\sqrt{n}}$

10. $\displaystyle\sum_{n=1}^{\infty} \frac{(iz)^n}{n^2}$

11. $\displaystyle\sum_{n=0}^{\infty} \frac{(n!)^3 z^n}{(3n)!}$

12. $\displaystyle\sum_{n=0}^{\infty} \frac{(n!)^2 z^n}{(2n)!}$

13. $\displaystyle\sum_{n=1}^{\infty} \frac{(z-i)^n}{n}$

14. $\displaystyle\sum_{n=0}^{\infty} n(n+1)(z-2i)^n$

15. $\displaystyle\sum_{n=0}^{\infty} \frac{(z-2+i)^n}{2^n}$

16. $\displaystyle\sum_{n=1}^{\infty} 2^n (z+i-3)^{2n}$

FEUE WHEP
10302

17. Verify the series in (7.3) by computer. Also show that it can be written in the form

$$\sum_{n=0}^{\infty}(-1)^n z^{2n} \sum_{k=0}^{n} \frac{1}{(2k+1)!}.$$

Use this form to show by ratio test that the series converges in the disk $|z| < 1$.

▶ 8. ELEMENTARY FUNCTIONS OF COMPLEX NUMBERS

The so-called elementary functions are powers and roots, trigonometric and inverse trigonometric functions, logarithmic and exponential functions, and combinations of these. All these you can compute or find in tables, as long as you want them as functions of real numbers. Now we want to find things like i^i, $\sin(1+i)$, or $\ln i$. These are not just curiosities for the amusement of the mathematically inclined, but may turn up to be evaluated in applied problems. To be sure, the values of experimental measurements are not imaginary. But the values of $\operatorname{Re} z$, $\operatorname{Im} z$, $|z|$, angle of z, are real, and these are the quantities which have experimental meaning. Meanwhile, mathematical solutions of problems may involve manipulations of complex numbers before we arrive finally at a real answer to compare with experiment.

Polynomials and rational functions (quotients of polynomials) of z are easily evaluated.

Example. If $f(z) = (z^2+1)/(z-3)$, we find $f(i-2)$ by substituting $z = i - 2$:

$$f(i-2) = \frac{(i-2)^2+1}{i-2-3} = \frac{-4i+4}{i-5} \cdot \frac{-i-5}{-i-5} = \frac{8i-12}{13}.$$

Next we want to investigate the possible meaning of other functions of complex numbers. We should like to define expressions like e^z or $\sin z$ so that they will obey the familiar laws we know for the corresponding real expressions [for example, $\sin 2x = 2\sin x \cos x$, or $(d/dx)e^x = e^x$]. We must, for consistency, define functions of complex numbers so that any equations involving them reduce to correct real equations when $z = x + iy$ becomes $z = x$, that is, when $y = 0$. These requirements will be met if we define e^z by the power series

$$e^z = \sum_{0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots. \tag{8.1}$$

This series converges for all values of the complex number z (Problem 7.1) and therefore gives us the value of e^z for any z. If we put $z = x$ (x real), we get the familiar series for e^x.

It is easy to show, by multiplying the series (Problem 1), that

$$e^{z_1} \cdot e^{z_2} = e^{z_1+z_2}. \tag{8.2}$$

In Chapter 14 we shall consider in detail the meaning of derivatives with respect to a complex z. However, it is worth while for you to know that $(d/dz)z^n = nz^{n-1}$, and that, in fact, the other differentiation and integration formulas which you know

FEUE WHEP

from elementary calculus hold also with x replaced by z. You can verify that $(d/dz)e^z = e^z$ when e^z is defined by (8.1) by differentiating (8.1) term by term (Problem 2). It can be shown that (8.1) is the only definition of e^z which preserves these familiar formulas. We now want to consider the consequences of this definition.

PROBLEMS, SECTION 8

Show from the power series (8.1) that

1. $e^{z_1} \cdot e^{z_2} = e^{z_1+z_2}$

2. $\dfrac{d}{dz}e^z = e^z$

3. Find the power series for $e^x \cos x$ and for $e^x \sin x$ from the series for e^z in the following way: Write the series for e^z; put $z = x + iy$. Show that $e^z = e^x(\cos y + i \sin y)$; take real and imaginary parts of the equation, and put $y = x$.

9. EULER'S FORMULA

For real θ, we know from Chapter 1 the power series for $\sin\theta$ and $\cos\theta$:

$$\begin{aligned} \sin\theta &= \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots, \\ \cos\theta &= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots. \end{aligned} \tag{9.1}$$

From our definition (8.1), we can write the series for e to any power, real or imaginary. We write the series for $e^{i\theta}$, where θ is real:

$$\begin{aligned} e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \cdots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} \cdots \\ &= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \cdots + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} \cdots\right). \end{aligned} \tag{9.2}$$

(The rearrangement of terms is justified because the series is absolutely convergent.) Now compare (9.1) and (9.2); the last line in (9.2) is just $\cos\theta + i\sin\theta$. We then have the very useful result we introduced in Section 3, known as Euler's formula:

$$e^{i\theta} = \cos\theta + i\sin\theta. \tag{9.3}$$

Thus we have justified writing any complex number as we did in (4.1), namely

$$z = x + iy = r(\cos\theta + i\sin\theta) = re^{i\theta}. \tag{9.4}$$

10304

Here are some examples of the use of (9.3) and (9.4). These problems can be done very quickly graphically or just by picturing them in your mind.

▶ **Examples.** Find the values of $2\,e^{i\pi/6}$, $e^{i\pi}$, $3\,e^{-i\pi/2}$, $e^{2n\pi i}$.

$2\,e^{i\pi/6}$ is $re^{i\theta}$ with $r = 2$, $\theta = \pi/6$. From Figure 9.1, $x = \sqrt{3}$, $y = 1$, $x + iy = \sqrt{3} + i$, so $2\,e^{i\pi/6} = \sqrt{3} + i$.

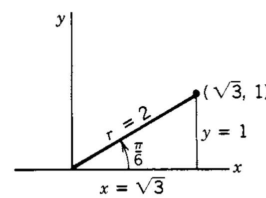


Figure 9.1

$e^{i\pi}$ is $re^{i\theta}$ with $r = 1$, $\theta = \pi$. From Figure 9.2, $x = -1$, $y = 0$, $x + iy = -1 + 0i$, so $e^{i\pi} = -1$. Note that $r = 1$ and $\theta = -\pi, \pm 3\pi, \pm 5\pi, \cdots$, give the same point, so $e^{-i\pi} = -1$, $e^{3\pi i} = -1$, and so on.

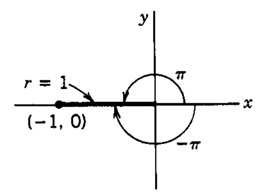


Figure 9.2

$3e^{-i\pi/2}$ is $re^{i\theta}$ with $r = 3$, $\theta = -\pi/2$. From Figure 9.3, $x = 0$, $y = -3$, so $3e^{-i\pi/2} = x + iy = 0 - 3i = -3i$.

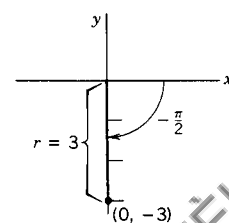


Figure 9.3

$e^{2n\pi i}$ is $re^{i\theta}$ with $r = 1$ and $\theta = 2n\pi = n(2\pi)$; that is, θ is an integral multiple of 2π. From Figure 9.4, $x = 1$, $y = 0$, so $e^{2n\pi i} = 1 + 0i = 1$.

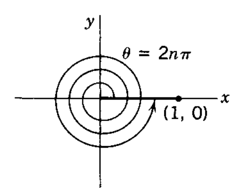


Figure 9.4

It is often convenient to use Euler's formula when we want to multiply or divide complex numbers. From (8.2) we obtain two familiar looking laws of exponents which are now valid for imaginary exponents:

$$\begin{aligned} e^{i\theta_1} \cdot e^{i\theta_2} &= e^{i(\theta_1+\theta_2)}, \\ e^{i\theta_1} \div e^{i\theta_2} &= e^{i(\theta_1-\theta_2)}. \end{aligned} \tag{9.5}$$

Remembering that *any* complex number can be written in the form $re^{i\theta}$ by (9.4), we get

FEUE WHEP

(9.6)
$$z_1 \cdot z_2 = r_1\, e^{i\theta_1} \cdot r_2\, e^{i\theta_2} = r_1\, r_2\, e^{i(\theta_1+\theta_2)},$$
$$z_1 \div z_2 = \frac{r_1}{r_2}\, e^{i(\theta_1-\theta_2)}.$$

In words, to multiply two complex numbers, we multiply their absolute values and add their angles. To divide two complex numbers, we divide the absolute values and subtract the angles.

Example. Evaluate $(1+i)^2/(1-i)$. From Figure 5.1 we have $1+i = \sqrt{2}\, e^{i\pi/4}$. We plot $1-i$ in Figure 9.5 and find $r = \sqrt{2}$, $\theta = -\pi/4$ (or $+7\pi/4$), so $1-i = \sqrt{2}\, e^{-i\pi/4}$. Then

$$\frac{(1+i)^2}{1-i} = \frac{(\sqrt{2}\, e^{i\pi/4})^2}{\sqrt{2}\, e^{-i\pi/4}} = \frac{2\, e^{i\pi/2}}{\sqrt{2}\, e^{-i\pi/4}} = \sqrt{2}\, e^{3i\pi/4}.$$

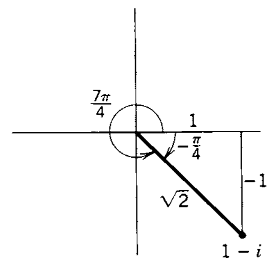


Figure 9.5

From Figure 9.6, we find $x = -1$, $y = 1$, so

$$\frac{(1+i)^2}{1-i} = x + iy = -1 + i.$$

We could use degrees in this problem. By (9.6), we find that the angle of $(1+i)^2/(1-i)$ is $2(45°) - (-45°) = 135°$ as in Figure 9.6.

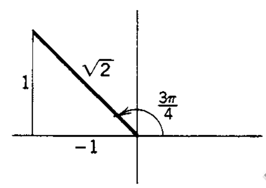


Figure 9.6

PROBLEMS, SECTION 9

Express the following complex numbers in the $x + iy$ form. Try to visualize each complex number, using sketches as in the examples if necessary. The first twelve problems you should be able to do in your head (and maybe some of the others—try it!) Doing a problem quickly in your head saves time over using a computer. Remember that the point in doing problems like this is to gain skill in manipulating complex expressions, so a good study method is to do the problems by hand and use a computer to check your answers.

1. $e^{-i\pi/4}$ **2.** $e^{i\pi/2}$ **3.** $9\, e^{3\pi i/2}$

4. $e^{(1/3)(3+4\pi i)}$ **5.** $e^{5\pi i}$ **6.** $e^{-2\pi i} - e^{-4\pi i} + e^{-6\pi i}$

7. $3\, e^{2(1+i\pi)}$ **8.** $2\, e^{5\pi i/6}$ **9.** $2\, e^{-i\pi/2}$

10. $e^{i\pi} + e^{-i\pi}$ **11.** $\sqrt{2}\, e^{5i\pi/4}$ **12.** $4\, e^{-8i\pi/3}$

13. $\dfrac{\left(i-\sqrt{3}\,\right)^3}{1-i}$ **14.** $\left(1+i\sqrt{3}\right)^6$ **15.** $(1+i)^2 + (1+i)^4$

16. $\left(i-\sqrt{3}\right)\left(1+i\sqrt{3}\right)$ **17.** $\dfrac{1}{(1+i)^3}$ **18.** $\left(\dfrac{1+i}{1-i}\right)^4$

19. $(1-i)^8$ **20.** $\left(\dfrac{\sqrt{2}}{i-1}\right)^{10}$ **21.** $\left(\dfrac{1-i}{\sqrt{2}}\right)^{40}$

10306

22. $\left(\dfrac{1-i}{\sqrt{2}}\right)^{42}$ **23.** $\dfrac{(1+i)^{48}}{(\sqrt{3}-i)^{25}}$ **24.** $\dfrac{(1-i\sqrt{3})^{21}}{(i-1)^{38}}$

25. $\left(\dfrac{i\sqrt{2}}{1+i}\right)^{12}$ **26.** $\left(\dfrac{2i}{i+\sqrt{3}}\right)^{19}$

27. Show that for any real y, $|e^{iy}|=1$. Hence show that $|e^z|=e^x$ for every complex z.

28. Show that the absolute value of a product of two complex numbers is equal to the product of the absolute values. Also show that the absolute value of the quotient of two complex numbers is the quotient of the absolute values. *Hint:* Write the numbers in the $re^{i\theta}$ form.

Use Problems 27 and 28 to find the following absolute values. If you understand Problems 27 and 28 and equation (5.1), you should be able to do these in your head.

29. $|e^{i\pi/2}|$ **30.** $|e^{\sqrt{3}-i}|$ **31.** $|5\,e^{2\pi i/3}|$ **32.** $|3e^{2+4i}|$

33. $|2\,e^{3+i\pi}|$ **34.** $|4\,e^{2i-1}|$ **35.** $|3\,e^{5i}\cdot 7\,e^{-2i}|$ **36.** $|2\,e^{i\pi/6}|^2$

37. $\left|\dfrac{1+i}{1-i}\right|$ **38.** $\left|\dfrac{e^{i\pi}}{1+i}\right|$

10. POWERS AND ROOTS OF COMPLEX NUMBERS

Using the rules (9.6) for multiplication and division of complex numbers, we have

$$z^n = (re^{i\theta})^n = r^n e^{in\theta} \tag{10.1}$$

for any integral n. In words, to obtain the nth power of a complex number, we take the nth power of the modulus and multiply the angle by n. The case $r=1$ is of particular interest. Then (10.1) becomes DeMoivre's theorem:

$$(e^{i\theta})^n = (\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta. \tag{10.2}$$

You can use this equation to find the formulas for $\sin 2\theta$, $\cos 2\theta$, $\sin 3\theta$, etc. (Problems 27 and 28).

The nth root of z, $z^{1/n}$, means a complex number whose nth power is z. From (10.1) you can see that this is

$$z^{1/n} = (re^{i\theta})^{1/n} = r^{1/n}e^{i\theta/n} = \sqrt[n]{r}\left(\cos\frac{\theta}{n} + i\sin\frac{\theta}{n}\right). \tag{10.3}$$

This formula must be used with care (see Examples 2 to 4 below).

Some examples will show how useful these formulas are.

Example 1.

$$[\cos(\pi/10) + i\sin(\pi/10)]^{25} = (e^{i\pi/10})^{25} = e^{2\pi i}e^{i\pi/2} = 1\cdot i = i.$$

FEUE WHEP

Example 2. Find the cube roots of 8. We know that 2 is a cube root of 8, but there are also two complex cube roots of 8; let us see why. Plot the complex number 8 (that is, $x = 8$, $y = 0$) in the complex plane; the polar coordinates of the point are $r = 8$, and $\theta = 0$, or $360°, 720°, 1080°$, etc. (We can use either degrees or radians here; read the end of Section 3.) Now by equation (10.3), $z^{1/3} = r^{1/3}e^{i\theta/3}$; that is, to find the polar coordinates of the cube root of a number $re^{i\theta}$, we find the cube root of r and divide the angle by 3. Then the polar coordinates of $\sqrt[3]{8}$ are

$$\begin{aligned} r = 2, \quad \theta &= 0°, \quad 360°/3, \quad 720°/3, \quad 1080°/3 \cdots \\ &= 0°, \quad 120°, \quad 240°, \quad 360° \cdots . \end{aligned} \tag{10.4}$$

We plot these points in Figure 10.1. Observe that the point $(2, 0°)$ and the point $(2, 360°)$ are the same. The points in (10.4) are all on a circle of radius 2 and are equally spaced $360°/3 = 120°$ apart. Starting with $\theta = 0$, if we add $120°$ repeatedly, we just repeat the three angles shown. Thus, there are exactly three cube roots for any number z, always on a circle of radius $\sqrt[3]{|z|}$ and spaced $120°$ apart.

Now to find the values of $\sqrt[3]{8}$ in rectangular form, we can read them from Figure 10.1, or we can calculate them from $z = r(\cos\theta + i\sin\theta)$ with $r = 2$ and $\theta = 0,\ 120° = 2\pi/3,\ 240° = 4\pi/3$. We can also use a computer to solve the equation $z^3 = 8$. By any of these methods we find

$$\sqrt[3]{8} = \{2,\ -1 + i\sqrt{3},\ -1 - i\sqrt{3}\}.$$

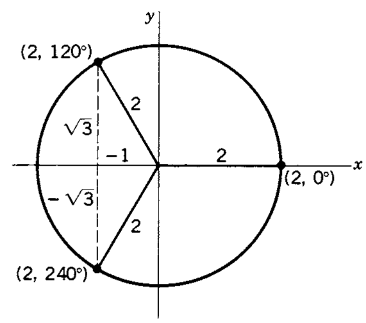


Figure 10.1

Example 3. Find and plot all values of $\sqrt[4]{-64}$. From Figure 10.2 (or by visualizing a plot of -64), we see that the polar coordinates of -64 are $r = 64$, $\theta = \pi + 2k\pi$

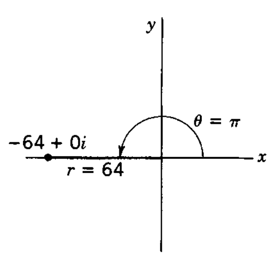


Figure 10.2

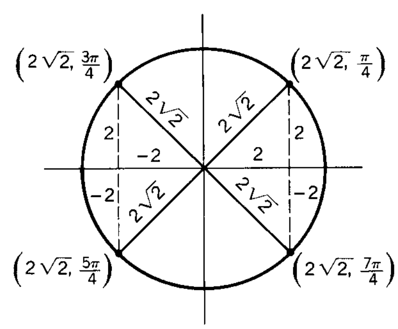


Figure 10.3

(where $k = 0, 1, 2, \cdots$). Then since $z^{1/4} = r^{1/4}e^{i\theta/4}$, the polar coordinates of $\sqrt[4]{-64}$ are

$$r = \sqrt[4]{64} = 2\sqrt{2},$$
$$\theta = \frac{\pi}{4}, \frac{\pi + 2\pi}{4}, \frac{\pi + 4\pi}{4}, \frac{\pi + 6\pi}{4}, \cdots = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}.$$

We plot these points in Figure 10.3. Observe that they are all on a circle of radius $2\sqrt{2}$, equally spaced $2\pi/4 = \pi/2$ apart. Starting with $\theta = \pi/4$, we add $\pi/2$

FEUE WHEP
10308

repeatedly, and find exactly 4 fourth roots. We can read the values of $\sqrt[4]{-64}$ in rectangular form from Figure 10.3:

$$\sqrt[4]{-64} = \pm 2 \pm 2i \quad \text{(all four combinations of } \pm \text{ signs)}$$

or we can calculate them as in Example 2, or we can solve the equation $z^4 = -64$ by computer.

▶ **Example 4.** Find and plot all values of $\sqrt[6]{-8i}$. The polar coordinates of $-8i$ are $r = 8$, $\theta = 270° + 360°k = 3\pi/2 + 2\pi k$. Then the polar coordinates of $\sqrt[6]{-8i}$ are

$$(10.5) \qquad r = \sqrt{2}, \quad \theta = \frac{270° + 360°k}{6} = 45° + 60°k \quad \text{or} \quad \theta = \frac{\pi}{4} + \frac{\pi}{3}k.$$

In Figure 10.4, we sketch a circle of radius $\sqrt{2}$. On it we plot the point at 45° and then plot the rest of the 6 equally spaced points 60° apart. To find the roots in rectangular coordinates, we need to find all the values of $r(\cos\theta + i\sin\theta)$ with r and θ given by (10.5). We can do this one root at a time or more simply by using a computer to solve the equation $z^6 = -8i$. We find (see Problem 33)

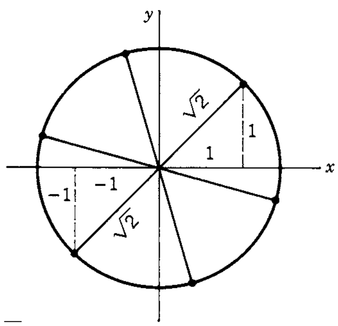


Figure 10.4

$$\pm\left\{1+i,\ \frac{\sqrt{3}+1}{2} - \frac{\sqrt{3}-1}{2}i,\ \frac{\sqrt{3}-1}{2} - \frac{\sqrt{3}+1}{2}i\right\} =$$
$$\pm\{1+i,\ 1.366 - 0.366i,\ 0.366 - 1.366i\}.$$

Summary In each of the preceding examples, our steps in finding $\sqrt[n]{re^{i\theta}}$ were:

(a) Find the polar coordinates of the roots: Take the nth root of r and divide $\theta + 2k\pi$ by n.

(b) Make a sketch: Draw a circle of radius $\sqrt[n]{r}$, plot the root with angle θ/n, and then plot the rest of the n roots around the circle equally spaced $2\pi/n$ apart. Note that we have now essentially solved the problem. From the sketch you can see the approximate rectangular coordinates of the roots and check your answers in (c). Since this sketch is quick and easy to do, it is worthwhile even if you use a computer to do part (c).

(c) Find the $x + iy$ coordinates of the roots by one of the methods in the examples. If you are using a computer, you may want to make a computer plot of the roots which should be a perfected copy of your sketch in (b).

▶ PROBLEMS, SECTION 10

Follow steps (a), (b), (c) above to find all the values of the indicated roots.

1.	$\sqrt[3]{1}$	**2.**	$\sqrt[3]{27}$	**3.**	$\sqrt[4]{1}$
4.	$\sqrt[4]{16}$	**5.**	$\sqrt[6]{1}$	**6.**	$\sqrt[6]{64}$
7.	$\sqrt[8]{16}$	**8.**	$\sqrt[8]{1}$	**9.**	$\sqrt[5]{1}$
10.	$\sqrt[5]{32}$	**11.**	$\sqrt[3]{-8}$	**12.**	$\sqrt[3]{-1}$

10309

13. $\sqrt[4]{-4}$ **14.** $\sqrt[4]{-1}$ **15.** $\sqrt[6]{-64}$

16. $\sqrt[6]{-1}$ **17.** $\sqrt[5]{-1}$ **18.** $\sqrt{i}$

19. $\sqrt[3]{i}$ **20.** $\sqrt[3]{-8i}$ **21.** $\sqrt{2+2i\sqrt{3}}$

22. $\sqrt[3]{2i-2}$ **23.** $\sqrt[4]{8i\sqrt{3}-8}$ **24.** $\sqrt[8]{\dfrac{-1-i\sqrt{3}}{2}}$

25. $\sqrt[5]{-1-i}$ **26.** $\sqrt[5]{i}$

27. Using the fact that a complex equation is really two real equations, find the double angle formulas (for $\sin 2\theta$, $\cos 2\theta$) by using equation (10.2).

28. As in Problem 27, find the formulas for $\sin 3\theta$ and $\cos 3\theta$.

29. Show that the center of mass of three identical particles situated at the points z_1, z_2, z_3 is $(z_1 + z_2 + z_3)/3$.

30. Show that the sum of the three cube roots of 8 is zero.

31. Show that the sum of the n nth roots of any complex number is zero.

32. The three cube roots of $+1$ are often called 1, ω, and ω^2. Show that this is reasonable, that is, show that the cube roots of $+1$ are $+1$ and two other numbers, each of which is the square of the other.

33. Verify the results given for the roots in Example 4. You can find the exact values in terms of $\sqrt{3}$ by using trigonometric addition formulas or more easily by using a computer to solve $z^6 = -8i$. (You still may have to do a little work by hand to put the computer's solution into the given form.)

11. THE EXPONENTIAL AND TRIGONOMETRIC FUNCTIONS

Although we have already defined e^z by a power series (8.1), it is worth while to write it in another form. By (8.2) we can write

$$e^z = e^{x+iy} = e^x e^{iy} = e^x(\cos y + i \sin y). \tag{11.1}$$

This is more convenient to use than the infinite series if we want values of e^z for given z. For example,

$$e^{2-i\pi} = e^2 e^{-i\pi} = e^2 \cdot (-1) = -e^2$$

from Figure 9.2.

We have already seen that there is a close relationship [Euler's formula (9.3)] between complex exponentials and trigonometric functions of real angles. It is useful to write this relation in another form. We write Euler's formula (9.3) as it

FEUE WHEP 10310

is, and also write it with θ replaced by $-\theta$. Remember that $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$. Then we have

$$(11.2)\qquad \begin{aligned} e^{i\theta} &= \cos\theta + i\sin\theta, \\ e^{-i\theta} &= \cos\theta - i\sin\theta. \end{aligned}$$

These two equations can be solved for $\sin\theta$ and $\cos\theta$. We get (Problem 2)

$$(11.3)\qquad \begin{aligned} \sin\theta &= \frac{e^{i\theta} - e^{-i\theta}}{2i}, \\ \cos\theta &= \frac{e^{i\theta} + e^{-i\theta}}{2}. \end{aligned}$$

These formulas are useful in evaluating integrals since products of exponentials are easier to integrate than products of sines and cosines. (See Problems 11 to 16, and Chapter 7, Section 5.)

So far we have discussed only trigonometric functions of real angles. We could define $\sin z$ and $\cos z$ for complex z by their power series as we did for e^z. We could then compare these series with the series for e^{iz} and derive Euler's formula and (11.3) with θ replaced by z. However, it is simpler to use the complex equations corresponding to (11.3) as our definitions for $\sin z$ and $\cos z$. We define

$$(11.4)\qquad \begin{aligned} \sin z &= \frac{e^{iz} - e^{-iz}}{2i}, \\ \cos z &= \frac{e^{iz} + e^{-iz}}{2}. \end{aligned}$$

The rest of the trigonometric functions of z are defined in the usual way in terms of these; for example, $\tan z = \sin z/\cos z$.

▶ **Example 1.** $\cos i = \dfrac{e^{i\cdot i} + e^{-i\cdot i}}{2} = \dfrac{e^{-1} + e}{2} = 1.543\cdots$. (We will see in Section 15 that this expression is called the hyperbolic cosine of 1.)

▶ **Example 2.**

$$\begin{aligned} \sin\left(\frac{\pi}{2} + i\ln 2\right) &= \frac{e^{i(\pi/2 + i\ln 2)} - e^{-i(\pi/2 + i\ln 2)}}{2i} \\ &= \frac{e^{i\pi/2}e^{-\ln 2} - e^{-i\pi/2}e^{\ln 2}}{2i} \qquad \text{by (8.2).} \end{aligned}$$

From Figures 5.2 and 9.3, $e^{i\pi/2} = i$, and $e^{-i\pi/2} = -i$. By the definition of $\ln x$ [or see equations (13.1) and (13.2)], $e^{\ln 2} = 2$, so $e^{-\ln 2} = 1/e^{\ln 2} = 1/2$. Then

$$\sin\left(\frac{\pi}{2} + i\ln 2\right) = \frac{(i)(1/2) - (-i)(2)}{2i} = \frac{5}{4}.$$

10311

Notice from both these examples that sines and cosines of complex numbers may be greater than 1. As we shall see (Section 15), although $|\sin x| \le 1$ and $|\cos x| \le 1$ for *real* x, when z is a complex number, $\sin z$ and $\cos z$ can have *any* value we like.

Using the definitions (11.4) of $\sin z$ and $\cos z$, you can show that the familiar trigonometric identities and calculus formulas hold when θ is replaced by z.

Example 3. Prove that $\sin^2 z + \cos^2 z = 1$.

$$\sin^2 z = \left(\frac{e^{iz} - e^{-iz}}{2i}\right)^2 = \frac{e^{2iz} - 2 + e^{-2iz}}{-4},$$

$$\cos^2 z = \left(\frac{e^{iz} + e^{-iz}}{2}\right)^2 = \frac{e^{2iz} + 2 + e^{-2iz}}{4},$$

$$\sin^2 z + \cos^2 z = \frac{2}{4} + \frac{2}{4} = 1.$$

Example 4. Using the definitions (11.4), verify that $(d/dz)\sin z = \cos z$.

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i},$$

$$\frac{d}{dz}\sin z = \frac{1}{2i}(ie^{iz} + ie^{-iz}) = \frac{e^{iz} + e^{-iz}}{2} = \cos z.$$

PROBLEMS, SECTION 11

1. Define $\sin z$ and $\cos z$ by their power series. Write the power series for e^{iz}. By comparing these series obtain the definition (11.4) of $\sin z$ and $\cos z$.
2. Solve the equations $e^{i\theta} = \cos\theta + i\sin\theta$, $e^{-i\theta} = \cos\theta - i\sin\theta$, for $\cos\theta$ and $\sin\theta$ and so obtain equations (11.3).

Find each of the following in rectangular form $x + iy$ and check your results by computer. Remember to save time by doing as much as you can in your head.

3. $e^{-(i\pi/4)+\ln 3}$
4. $e^{3\ln 2 - i\pi}$
5. $e^{(i\pi/4)+(\ln 2)/2}$
6. $\cos(i\ln 5)$
7. $\tan(i\ln 2)$
8. $\cos(\pi - 2i\ln 3)$
9. $\sin(\pi - i\ln 3)$
10. $\sin(i\ln i)$

In the following integrals express the sines and cosines in exponential form and then integrate to show that:

11. $\displaystyle\int_{-\pi}^{\pi} \cos 2x\, \cos 3x\, dx = 0$
12. $\displaystyle\int_{-\pi}^{\pi} \cos^2 3x\, dx = \pi$
13. $\displaystyle\int_{-\pi}^{\pi} \sin 2x\, \sin 3x\, dx = 0$
14. $\displaystyle\int_{0}^{2\pi} \sin^2 4x\, dx = \pi$
15. $\displaystyle\int_{-\pi}^{\pi} \sin 2x\, \cos 3x\, dx = 0$
16. $\displaystyle\int_{-\pi}^{\pi} \sin 3x\, \cos 4x\, dx = 0$

Evaluate $\int e^{(a+ib)x}dx$ and take real and imaginary parts to show that:

17. $\displaystyle\int e^{ax}\cos bx\, dx = \frac{e^{ax}(a\cos bx + b\sin bx)}{a^2+b^2}$
18. $\displaystyle\int e^{ax}\sin bx\, dx = \frac{e^{ax}(a\sin bx - b\cos bx)}{a^2+b^2}$

10312

12. HYPERBOLIC FUNCTIONS

Let us look at $\sin z$ and $\cos z$ for pure imaginary z, that is, $z = iy$:

$$\begin{aligned} \sin iy &= \frac{e^{-y} - e^{y}}{2i} = i\,\frac{e^{y} - e^{-y}}{2}, \\ \cos iy &= \frac{e^{-y} + e^{y}}{2} = \frac{e^{y} + e^{-y}}{2}. \end{aligned} \tag{12.1}$$

The real functions on the right have special names because these particular combinations of exponentials arise frequently in problems. They are called the hyperbolic sine (abbreviated sinh) and the hyperbolic cosine (abbreviated cosh). Their definitions for all z are

$$\begin{aligned} \sinh z &= \frac{e^{z} - e^{-z}}{2}, \\ \cosh z &= \frac{e^{z} + e^{-z}}{2}. \end{aligned} \tag{12.2}$$

The other hyperbolic functions are named and defined in a similar way to parallel the trigonometric functions:

$$\begin{aligned} \tanh z &= \frac{\sinh z}{\cosh z}, & \coth z &= \frac{1}{\tanh z}, \\ \operatorname{sech} z &= \frac{1}{\cosh z}, & \operatorname{csch} z &= \frac{1}{\sinh z}. \end{aligned} \tag{12.3}$$

(See Problem 38 for the reason behind the term "hyperbolic" functions.)

We can write (12.1) as

$$\begin{aligned} \sin iy &= i\,\sinh y, \\ \cos iy &= \cosh y. \end{aligned} \tag{12.4}$$

Then we see that the hyperbolic functions of y are (except for one i factor) the trigonometric functions of iy. From (12.2) we can show that (12.4) holds with y replaced by z. Because of this relation between hyperbolic and trigonometric functions, the formulas for hyperbolic functions look very much like the corresponding trigonometric identities and calculus formulas. They are not identical, however.

▶ **Example.** You can prove the following formulas (see Problems 9, 10, 11 and 38).

$$\begin{aligned} &\cosh^2 z - \sinh^2 z = 1 && \left(\text{compare } \sin^2 z + \cos^2 z = 1\right), \\ &\frac{d}{dz}\cosh z = \sinh z && \left(\text{compare } \frac{d}{dz}\cos z = -\sin z\right). \end{aligned}$$

PROBLEMS, SECTION 12

Verify each of the following by using equations (11.4), (12.2), and (12.3).

1. $\sin z = \sin(x + iy) = \sin x\, \cosh y + i\, \cos x\, \sinh y$

FEUE WHEP

2. $\cos z = \cos x \cosh y - i \sin x \sinh y$

3. $\sinh z = \sinh x \cos y + i \cosh x \sin y$

4. $\cosh z = \cosh x \cos y + i \sinh x \sin y$

5. $\sin 2z = 2 \sin z \cos z$

6. $\cos 2z = \cos^2 z - \sin^2 z$

7. $\sinh 2z = 2 \sinh z \cosh z$

8. $\cosh 2z = \cosh^2 z + \sinh^2 z$

9. $\dfrac{d}{dz} \cos z = -\sin z$

10. $\dfrac{d}{dz} \cosh z = \sinh z$

11. $\cosh^2 z - \sinh^2 z = 1$

12. $\cos^4 z + \sin^4 z = 1 - \dfrac{1}{2} \sin^2 2z$

13. $\cos 3z = 4 \cos^3 z - 3 \cos z$

14. $\sin iz = i \sinh z$

15. $\sinh iz = i \sin z$

16. $\tan iz = i \tanh z$

17. $\tanh iz = i \tan z$

18. $\tan z = \tan(x + iy) = \dfrac{\tan x + i \tanh y}{1 - i \tan x \tanh y}$

19. $\tanh z = \dfrac{\tanh x + i \tan y}{1 + i \tanh x \tan y}$

20. Show that $e^{nz} = (\cosh z + \sinh z)^n = \cosh nz + \sinh nz$. Use this and a similar equation for e^{-nz} to find formulas for $\cosh 3z$ and $\sinh 3z$ in terms of $\sinh z$ and $\cosh z$.

21. Use a computer to plot graphs of $\sinh x$, $\cosh x$, and $\tanh x$.

22. Using (12.2) and (8.1), find, in summation form, the power series for $\sinh x$ and $\cosh x$. Check the first few terms of your series by computer.

Find the real part, the imaginary part, and the absolute value of

23. $\cosh(ix)$

24. $\cos(ix)$

25. $\sin(x - iy)$

26. $\cosh(2 - 3i)$

27. $\sin(4 + 3i)$

28. $\tanh(1 - i\pi)$

Find each of the following in the $x + iy$ form and check your answers by computer.

29. $\cosh 2\pi i$

30. $\tanh \dfrac{3\pi i}{4}$

31. $\sinh \left(\ln 2 + \dfrac{i\pi}{3} \right)$

32. $\cosh \left(\dfrac{i\pi}{2} - \ln 3 \right)$

33. $\tan i$

34. $\sin \dfrac{i\pi}{2}$

35. $\cosh(i\pi + 2)$

36. $\sinh \left(1 + \dfrac{i\pi}{2} \right)$

37. $\cos(i\pi)$

38. The functions $\sin t$, $\cos t$, $\cdots$, are called "circular functions" and the functions $\sinh t$, $\cosh t$, $\cdots$, are called "hyperbolic functions". To see a reason for this, show that $x = \cos t$, $y = \sin t$, satisfy the equation of a circle $x^2 + y^2 = 1$, while $x = \cosh t$, $y = \sinh t$, satisfy the equation of a hyperbola $x^2 - y^2 = 1$.

10314

▶ 13. LOGARITHMS

In elementary mathematics you learned to find logarithms of positive numbers only; in fact, you may have been told that there were no logarithms of negative numbers. This is true if you use only real numbers, but it is not true when we allow complex numbers as answers. We shall now see how to find the logarithm of any complex number $z \neq 0$ (including negative real numbers as a special case). If

$$z = e^w, \tag{13.1}$$

then by definition

$$w = \ln z. \tag{13.2}$$

(We use ln for natural logarithms to avoid the cumbersome $\log_e$ and to avoid confusion with logarithms to the base 10.)

We can write the law of exponents (8.2), using the letters of (13.1), as

$$z_1 z_2 = e^{w_1} \cdot e^{w_2} = e^{w_1+w_2}. \tag{13.3}$$

Taking logarithms of this equation, that is, using (13.1) and (13.2), we get

$$\ln z_1 z_2 = w_1 + w_2 = \ln z_1 + \ln z_2. \tag{13.4}$$

This is the familiar law for the logarithm of a product, justified now for complex numbers. We can then find the real and imaginary parts of the logarithm of a complex number from the equation

$$w = \ln z = \ln(re^{i\theta}) = \operatorname{Ln} r + \ln e^{i\theta} = \operatorname{Ln} r + i\theta, \tag{13.5}$$

where $\operatorname{Ln} r$ means the ordinary real logarithm to the base e of the real positive number r.

Since θ has an infinite number of values (all differing by multiples of 2π), a complex number has infinitely many logarithms, differing from each other by multiples of $2\pi i$. The *principal value* of $\ln z$ (often written as $\operatorname{Ln} z$) is the one using the principal value of θ, that is $0 \le \theta < 2\pi$. (Some references use $-\pi < \theta \le \pi$.)

▶ **Example 1.** Find $\ln(-1)$. From Figure 9.2, we see that the polar coordinates of the point $z = -1$ are $r = 1$ and $\theta = \pi, -\pi, 3\pi, \cdots$. Then,

$$\ln(-1) = \operatorname{Ln}(1) + i(\pi \pm 2n\pi) = i\pi, -i\pi, 3\pi i, \cdots.$$

▶ **Example 2.** Find $\ln(1+i)$. From Figure 5.1, for $z = 1 + i$, we find $r = \sqrt{2}$, and $\theta = \pi/4 \pm 2n\pi$. Then

$$\ln(1+i) = \operatorname{Ln}\sqrt{2} + i\left(\frac{\pi}{4} \pm 2n\pi\right) = 0.347\cdots + i\left(\frac{\pi}{4} \pm 2n\pi\right).$$

Even a positive real number now has infinitely many logarithms, since its angle can be taken as 0, 2π, -2π, etc. Only one of these logarithms is real, namely the principal value $\operatorname{Ln} r$ using the angle $\theta = 0$.

10315

14. COMPLEX ROOTS AND POWERS

For real positive numbers, the equation $\ln a^b = b \ln a$ is equivalent to $a^b = e^{b \ln a}$. We define complex powers by the same formula with complex a and b. By definition, for complex a and b ($a \neq e$),

$$a^b = e^{b \ln a}. \tag{14.1}$$

[The case $a = e$ is excluded because we have already defined powers of e by (8.1).] Since $\ln a$ is multiple valued (because of the infinite number of values of θ), powers a^b are usually multiple valued, and unless you want just the principal value of $\ln z$ or of a^b you must use all values of θ. In the following examples we find all values of each complex power and write the answers in the $x + iy$ form.

Example 1. Find all values of i^{-2i}. From Figure 5.2, and equation (13.5) we find $\ln i = \text{Ln}\, 1 + i(\pi/2 \pm 2n\pi) = i(\pi/2 \pm 2n\pi)$ since $\text{Ln}\, 1 = 0$. Then, by equation (14.1),

$$i^{-2i} = e^{-2i \ln i} = e^{-2i \cdot i(\pi/2 \pm 2n\pi)} = e^{\pi \pm 4n\pi} = e^{\pi}, e^{5\pi}, e^{-3\pi}, \cdots,$$

where $e^{\pi} = 23.14\cdots$. Note the infinite set of values of i^{-2i}, all real! Also read the end of Section 3, and note that here the final step is *not* to find sine or cosine of $\pi \pm 4n\pi$; thus, in finding $\ln i = i\theta$, we *must not* write θ in degrees.

Example 2. Find all values of $i^{1/2}$. Using $\ln i$ from Example 1 we have $i^{1/2} = e^{(1/2)\ln i} = e^{i(\pi/4 + n\pi)} = e^{i\pi/4} e^{in\pi}$. Now $e^{in\pi} = +1$ when n is even (Fig. 9.4), and $e^{in\pi} = -1$ when n is odd (Fig. 9.2). Thus,

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

using Figure 5.1. Notice that although $\ln i$ has an infinite set of values, we find just two values for $i^{1/2}$ as we should for a square root. (Compare the method of Section 10 which is easier for this problem.)

Example 3. Find all values of $(1+i)^{1-i}$. Using (14.1) and the value of $\ln(1+i)$ from Example 2, Section 13, we have

$$\begin{aligned}
(1+i)^{1-i} &= e^{(1-i)\ln(1+i)} = e^{(1-i)[\text{Ln}\sqrt{2} + i(\pi/4 \pm 2n\pi)]} \\
&= e^{\text{Ln}\sqrt{2}}\, e^{-i\,\text{Ln}\sqrt{2}}\, e^{i\pi/4}\, e^{\pm 2n\pi i}\, e^{\pi/4}\, e^{\pm 2n\pi} \\
&= \sqrt{2}\, e^{i(\pi/4 - \text{Ln}\sqrt{2})}\, e^{\pi/4}\, e^{\pm 2n\pi} \qquad (\text{since } e^{\pm 2n\pi i} = 1) \\
&= \sqrt{2}\, e^{\pi/4}\, e^{\pm 2n\pi}[\cos(\pi/4 - \text{Ln}\sqrt{2}) + i\sin(\pi/4 - \text{Ln}\sqrt{2})] \\
&\cong e^{\pm 2n\pi}(2.808 + 1.318i).
\end{aligned}$$

Now you may be wondering why not just do these problems by computer. The most important point is that it is useful for advanced work to have skill in manipulating complex expressions. A second point is that there may be several forms for an answer (see Section 15, Example 2) or there may be many answers (see examples

10316

above), and your computer may not give you the one you want (see Problem 25). So to obtain needed skills, a good study method is to do problems by hand and compare with computer solutions.

▶ PROBLEMS, SECTION 14

Evaluate each of the following in $x + iy$ form, and compare with a computer solution.

1. $\ln(-e)$ **2.** $\ln(-i)$ **3.** $\ln(i+\sqrt{3})$

4. $\ln(i-1)$ **5.** $\ln(-\sqrt{2}-i\sqrt{2})$ **6.** $\ln\left(\dfrac{1-i}{\sqrt{2}}\right)$

7. $\ln\left(\dfrac{1+i}{1-i}\right)$ **8.** $i^{2/3}$ **9.** $(-1)^i$

10. $i^{\ln i}$ **11.** 2^i **12.** i^{3+i}

13. $i^{2i/\pi}$ **14.** $(2i)^{1+i}$ **15.** $(-1)^{\sin i}$

16. $\left(\dfrac{1+i\sqrt{3}}{2}\right)^i$ **17.** $(i-1)^{i+1}$ **18.** $\cos(2i\ln i)$

19. $\cos(\pi + i\ln 2)$ **20.** $\sin\left(i\ln\dfrac{1-i}{1+i}\right)$ **21.** $\cos[i\ln(-1)]$

22. $\sin\left[i\ln\left(\dfrac{\sqrt{3}+i}{2}\right)\right]$ **23.** $\left(1-\sqrt{2i}\right)^i$. *Hint:* Find $\sqrt{2i}$ first.

24. Show that $(a^b)^c$ can have more values than a^{bc}. As examples compare

(a) $[(-i)^{2+i}]^{2-i}$ and $(-i)^{(2+i)(2-i)} = (-i)^5$;

(b) $(i^i)^i$ and i^{-1}.

25. Use a computer to find the three solutions of the equation $x^3-3x-1=0$. Find a way to show that the solutions can be written as $2\cos(\pi/9)$, $-2\cos(2\pi/9)$, $-2\cos(4\pi/9)$.

▶ 15. INVERSE TRIGONOMETRIC AND HYPERBOLIC FUNCTIONS

We have already defined the trigonometric and hyperbolic functions of a complex number z. For example,

$$w = \cos z = \frac{e^{iz}+e^{-iz}}{2} \tag{15.1}$$

defines $w = \cos z$; that is, for each complex number z, (15.1) gives us the complex number w. We now define the inverse cosine or arc cos w by

$$z = \operatorname{arc\,cos} w \quad \text{if} \quad w = \cos z. \tag{15.2}$$

The other inverse trigonometric and hyperbolic functions are defined similarly.

In dealing with real numbers, you know that $\sin x$ and $\cos x$ are never greater than 1. This is no longer true for $\sin z$ and $\cos z$ with z complex. To illustrate the method of finding inverse trigonometric (or inverse hyperbolic) functions, let's find arc cos 2.

FEUE WHEP

Example 1. We want z, where

$$z = \arccos 2 \quad \text{or} \quad \cos z = 2.$$

Then we have

$$\frac{e^{iz} + e^{-iz}}{2} = 2.$$

To simplify the algebra, let $u = e^{iz}$. Then $e^{-iz} = u^{-1}$, and the equation becomes

$$\frac{u + u^{-1}}{2} = 2.$$

Multiply by 2 and by u to get $u^2 + 1 = 4u$ or $u^2 - 4u + 1 = 0$. Solve this equation by the quadratic formula to find

$$u = \frac{4 \pm \sqrt{16 - 4}}{2} = 2 \pm \sqrt{3}, \quad \text{or} \quad e^{iz} = u = 2 \pm \sqrt{3}.$$

Take logarithms of both sides of this equation, and solve for z:

$$iz = \ln(2 \pm \sqrt{3}) = \mathrm{Ln}(2 \pm \sqrt{3}) + 2n\pi i,$$
$$\arccos 2 = z = 2n\pi - i\,\mathrm{Ln}(2 \pm \sqrt{3}\,) = 2n\pi \pm i\,\mathrm{Ln}(2 + \sqrt{3}\,)$$

since $\mathrm{Ln}(2 - \sqrt{3}) = -\mathrm{Ln}(2 + \sqrt{3})$.

It is instructive now to find $\cos z$ and see that it *is* 2. For $iz = \ln(2 \pm \sqrt{3})$, we have

$$e^{iz} = e^{\ln(2 \pm \sqrt{3})} = 2 \pm \sqrt{3},$$
$$e^{-iz} = \frac{1}{e^{iz}} = \frac{1}{2 \pm \sqrt{3}} = \frac{2 \mp \sqrt{3}}{4 - 3} = 2 \mp \sqrt{3}.$$

Then

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} = \frac{2 \pm \sqrt{3} + 2 \mp \sqrt{3}}{2} = \frac{4}{2} = 2,$$

as claimed.

By the same method, we can find all the inverse trigonometric and hyperbolic functions in terms of logarithms. (See Problems, Section 17.) Here is one more example.

Example 2. In integral tables or from your computer you may find for the indefinite integral

$$\int \frac{dx}{\sqrt{x^2 + a^2}} \tag{15.3}$$

either

$$\sinh^{-1} \frac{x}{a} \quad \text{or} \quad \ln(x + \sqrt{x^2 + a^2}\,). \tag{15.4}$$

How are these related? Put

$$z = \sinh^{-1} \frac{x}{a} \quad \text{or} \quad \frac{x}{a} = \sinh z = \frac{e^z - e^{-z}}{2}. \tag{15.5}$$

FÊUE WHEP

10318

We solve for z as in the previous example. Let $e^z = u$, $e^{-z} = 1/u$. Then

$$u - \frac{1}{u} = \frac{2x}{a},$$
$$au^2 - 2xu - a = 0,$$

(15.6) $$e^z = u = \frac{2x \pm \sqrt{4x^2 + 4a^2}}{2a} = \frac{x \pm \sqrt{x^2 + a^2}}{a}.$$

For real integrals, that is, for real z, $e^z > 0$, so we must use the positive sign. Then, taking the logarithm of (15.6) we have

(15.7) $$z = \ln(x + \sqrt{x^2 + a^2}\,) - \ln a.$$

Comparing (15.5) and (15.7) we see that the two answers in (15.4) differ only by the constant $\ln a$, which is a constant of integration.

▶ PROBLEMS, SECTION 15

Find each of the following in the $x + iy$ form and compare a computer solution.

1. $\arcsin 2$

2. $\arctan 2i$

3. $\cosh^{-1}(1/2)$

4. $\sinh^{-1}(i/2)$

5. $\arccos\left(i\sqrt{8}\right)$

6. $\tanh^{-1}(-i)$

7. $\arctan\left(i\sqrt{2}\right)$

8. $\arcsin(5/3)$

9. $\tanh^{-1}\left(i\sqrt{3}\right)$

10. $\arccos(5/4)$

11. $\sinh^{-1}\left(i/\sqrt{2}\right)$

12. $\cosh^{-1}\left(\sqrt{3}/2\right)$

13. $\cosh^{-1}(-1)$

14. $\arcsin(3i/4)$

15. $\arctan(2 + i)$

16. $\tanh^{-1}(1 - 2i)$

17. Show that $\tan z$ never takes the values $\pm i$. *Hint:* Try to solve the equation $\tan z = i$ and find that it leads to a contradiction.

18. Show that $\tanh z$ never takes the values ± 1.

▶ 16. SOME APPLICATIONS

Motion of a Particle We have already seen (end of Section 5) that the path of a particle in the (x, y) plane is given by $z = z(t)$. As another example of this, suppose $z = 1 + 3e^{2it}$. We see that

(16.1) $$|z - 1| = |3\,e^{2it}| = 3.$$

Recall that $|z - 1|$ is the distance between the points z and 1; (16.1) says that this distance is 3. Thus the particle traverses a circle of radius 3, with center at $(1, 0)$. The magnitude of its velocity is $|dz/dt| = |6i\,e^{2it}| = 6$, so it moves around the circle at constant speed. (Also see Problem 2).

FEUE WHEP

PROBLEMS, SECTION 16

1. Show that if the line through the origin and the point z is rotated 90° about the origin, it becomes the line through the origin and the point iz. This fact is sometimes expressed by saying that multiplying a complex number by i rotates it through 90°. Use this idea in the following problem. Let $z = a\,e^{i\omega t}$ be the displacement of a particle from the origin at time t. Show that the particle travels in a circle of radius a at velocity $v = a\omega$ and with acceleration of magnitude v^2/a directed toward the center of the circle.

In each of the following problems, z represents the displacement of a particle from the origin. Find (as functions of t) its speed and the magnitude of its acceleration, and describe the motion.

2. $z = 5e^{i\omega t}$, $\omega = \text{const.}$ *Hint:* See Problem 1.

3. $z = (1+i)e^{it}$.

4. $z = (1+i)t - (2+i)(1-t)$. *Hint:* Show that the particle moves along a straight line through the points $(1+i)$ and $(-2-i)$.

5. $z = z_1 t + z_2(1-t)$. *Hint:* See Problem 4; the straight line here is through the points z_1 and z_2.

Electricity In the theory of electric circuits, it is shown that if V_R is the voltage across a resistance R, and I is the current flowing through the resistor, then

(16.2) $$V_R = IR \quad \text{(Ohm's law)}.$$

It is also known that the current and voltage across an inductance L are related by

(16.3) $$V_L = L\frac{dI}{dt}$$

and the current and voltage across a capacitor are related by

(16.4) $$\frac{dV_C}{dt} = \frac{I}{C},$$

where C is the capacitance. Suppose the current I and voltage V in the circuit of Figure 16.1 vary with time so that I is given by

(16.5) $$I = I_0 \sin\omega t.$$

You can verify that the following voltages across R, L, and C are consistent with (16.2), (16.3), and (16.4):

(16.6) $$V_R = RI_0 \sin\omega t,$$

(16.7) $$V_L = \omega L I_0 \cos\omega t,$$

(16.8) $$V_C = -\frac{1}{\omega C} I_0 \cos\omega t.$$

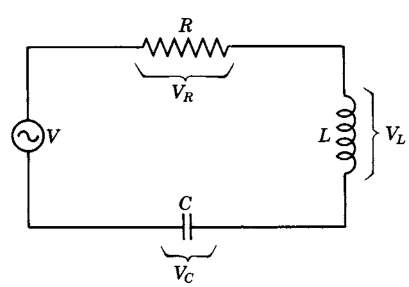


Figure 16.1

The total voltage

(16.9) $$V = V_R + V_L + V_C$$

10320

is then a complicated function. A simpler method of discussing a-c circuits uses complex quantities as follows. Instead of (16.5) we write

$$I = I_0\, e^{i\omega t}, \tag{16.10}$$

where it is understood that the actual physical current is given by the imaginary part of I in (16.10), that is, by (16.5). Note, by comparing (16.5) and (16.10), that the maximum value of I, namely I_0, is given in (16.10) by $|I|$. Now equations (16.6) to (16.9) become

$$V_R = RI_0\, e^{i\omega t} = RI, \tag{16.11}$$

$$V_L = i\omega LI_0\, e^{i\omega t} = i\omega LI, \tag{16.12}$$

$$V_C = \frac{1}{i\omega C} I_0\, e^{i\omega t} = \frac{1}{i\omega C} I, \tag{16.13}$$

$$V = V_R + V_L + V_C = \left[R + i\left(\omega L - \frac{1}{\omega C}\right)\right] I. \tag{16.14}$$

The complex quantity Z defined by

$$Z = R + i\left(\omega L - \frac{1}{\omega C}\right) \tag{16.15}$$

is called the (complex) impedance. Using it we can write (16.14) as

$$V = ZI \tag{16.16}$$

which looks much like Ohm's law. In fact, Z for an a-c circuit corresponds to R for a d-c circuit. The more complicated a-c circuit equations now take the same simple form as the d-c equations except that all quantities are complex. For example, the rules for combining resistances in series and in parallel hold for combining complex impedances (see Problems below).

▶ PROBLEMS, SECTION 16

In electricity we learn that the resistance of two resistors in series is $R_1 + R_2$ and the resistance of two resistors in parallel is $(R_1^{-1} + R_2^{-1})^{-1}$. Corresponding formulas hold for complex impedances. Find the impedance of Z_1 and Z_2 in series, and in parallel, given:

6. (a) $Z_1 = 2 + 3i$, $Z_2 = 1 - 5i$ (b) $Z_1 = 2\sqrt{3}\, e^{i\pi/6}$, $Z_2 = 2\, e^{2i\pi/3}$

7. (a) $Z_1 = 1 - i$, $Z_2 = 3i$ (b) $|Z_1| = 3.16$, $\theta_1 = 18.4°$; $|Z_2| = 4.47$, $\theta_2 = 63.4°$

8. Find the impedance of the circuit in Figure 16.2 (R and L in series, and then C in parallel with them). A circuit is said to be *in resonance* if Z is real; find ω in terms of R, L, and C at resonance.

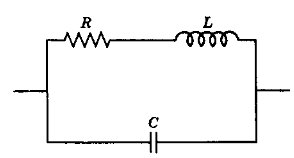


Figure 16.2

9. For the circuit in Figure 16.1:

(a) Find ω in terms of R, L, and C if the angle of Z is 45°.

(b) Find the resonant frequency ω (see Problem 8).

10. Repeat Problem 9 for a circuit consisting of R, L, and C, all in parallel.

FEUE WHEP
10321

Optics In optics we frequently need to combine a number of light waves (which can be represented by sine functions). Often each wave is "out of phase" with the preceding one by a fixed amount; this means that the waves can be written as $\sin t$, $\sin(t+\delta)$, $\sin(t+2\delta)$, and so on. Suppose we want to add all these sine functions together. An easy way to do it is to see that each sine is the imaginary part of a complex number, so what we want is the imaginary part of the series

$$e^{it} + e^{i(t+\delta)} + e^{i(t+2\delta)} + \cdots . \tag{16.17}$$

This is a geometric progression with first term e^{it} and ratio $e^{i\delta}$. If there are n waves to be combined, we want the sum of n terms of this progression, which is

$$\frac{e^{it}(1-e^{in\delta})}{1-e^{i\delta}}. \tag{16.18}$$

We can simplify this expression by writing

$$1 - e^{i\delta} = e^{i\delta/2}(e^{-i\delta/2} - e^{i\delta/2}) = -e^{i\delta/2} \cdot 2i \sin\frac{\delta}{2} \tag{16.19}$$

by (11.3). Substituting (16.19) and a similar formula for $(1-e^{in\delta})$ into (16.18), we get

$$\frac{e^{it}e^{in\delta/2}}{e^{i\delta/2}} \frac{\sin(n\delta/2)}{\sin(\delta/2)} = e^{i\{t+[(n-1)/2]\delta\}} \frac{\sin(n\delta/2)}{\sin(\delta/2)}. \tag{16.20}$$

The imaginary part of the series (16.17) which we wanted is then the imaginary part of (16.20), namely

$$\sin\left(t + \frac{n-1}{2}\delta\right) \sin\frac{n\delta}{2} \Big/ \sin\frac{\delta}{2}.$$

PROBLEMS, SECTION 16

11. Prove that

$$\cos\theta + \cos 3\theta + \cos 5\theta + \cdots + \cos(2n-1)\theta = \frac{\sin 2n\theta}{2\sin\theta},$$
$$\sin\theta + \sin 3\theta + \sin 5\theta + \cdots + \sin(2n-1)\theta = \frac{\sin^2 n\theta}{\sin\theta}.$$

Hint: Use Euler's formula and the geometric progression formula.

12. In optics, the following expression needs to be evaluated in calculating the intensity of light transmitted through a film after multiple reflections at the surfaces of the film:

$$\left(\sum_{n=0}^{\infty} r^{2n}\cos n\theta\right)^2 + \left(\sum_{n=0}^{\infty} r^{2n}\sin n\theta\right)^2.$$

Show that this is equal to $|\sum_{n=0}^{\infty} r^{2n}e^{in\theta}|^2$ and so evaluate it assuming $|r| < 1$ (r is the fraction of light reflected each time).

FEUE WHEP

Simple Harmonic Motion It is very convenient to use complex notation even for motion along a straight line. Think of a mass m attached to a spring and oscillating up and down (see Figure 16.3). Let y be the vertical displacement of the mass from its equilibrium position (the point at which it would hang at rest). Recall that the force on m due to the stretched or compressed spring is then $-ky$, where k is the spring constant, and the minus sign indicates that the force and displacement are in opposite directions. Then Newton's second law (force = mass times acceleration) gives

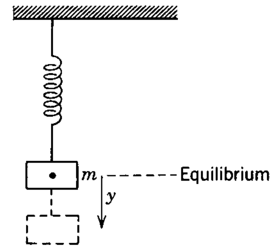


Figure 16.3

$$m\frac{d^2y}{dt^2} = -ky \quad \text{or} \quad \frac{d^2y}{dt^2} = -\frac{k}{m}y = -\omega^2 y \quad \text{if} \quad \omega^2 = \frac{k}{m}. \tag{16.21}$$

Now we want a function $y(t)$ with the property that differentiating it twice just multiplies it by a constant. You can easily verify that this is true for exponentials, sines, and cosines (see problem 13). Just as in discussing electric circuits (see (16.10)), we may write a solution of (16.21) as

$$y = y_0\, e^{i\omega t} \tag{16.22}$$

with the understanding that the actual physical displacement is either the real or the imaginary part of (16.22). The constant $\omega = \sqrt{k/m}$ is called the angular frequency (see Chapter 7, Section 2). We will use this notation in Chapter 3, Section 12.

▶ PROBLEMS, SECTION 16

13. Verify that $e^{i\omega t}$, $e^{-i\omega t}$, $\cos\omega t$, and $\sin\omega t$ satisfy equation (16.21).

▶ 17. MISCELLANEOUS PROBLEMS

Find one or more values of each of the following complex expressions and compare with a computer solution.

1. $\left(\dfrac{1+i}{1-i}\right)^{2718}$ **2.** $\left(\dfrac{1+i\sqrt{3}}{\sqrt{2}+i\sqrt{2}}\right)^{50}$ **3.** $\sqrt[5]{-4-4i}$

4. $\sinh(1+i\pi/2)$ **5.** $\tanh(i\pi/4)$ **6.** $(-e)^{i\pi}$

7. $(-i)^i$ **8.** $\cos\left[2i\ln\dfrac{1-i}{1+i}\right]$ **9.** $\arcsin\left[\left(\dfrac{\sqrt{3}+i}{\sqrt{3}-i}\right)^{12}\right]$

10. $e^{2i\arctan(i\sqrt{3})}$ **11.** $e^{2\tanh^{-1} i}$ **12.** $e^{i\arcsin i}$

13. Find *real* x and y for which $|z+3| = 1 - iz$, where $z = x+iy$.

14. Find the disk of convergence of the series $\sum(z-2i)^n/n$.

15. For what z is the series $\sum z^{\ln n}$ absolutely convergent? *Hints:* Use equation (14.1). Also see Chapter 1, Problem 6.15.

16. Describe the set of points z for which $\mathrm{Re}(e^{i\pi/2}z) > 2$.

FEUE WHEP

Verify the formulas in Problems 17 to 24.

17. $\arcsin z = -i\ln(iz \pm \sqrt{1-z^2})$

18. $\arccos z = i\ln(z \pm \sqrt{z^2-1})$

19. $\arctan z = \dfrac{1}{2i}\ln\dfrac{1+iz}{1-iz}$

20. $\sinh^{-1} z = \ln(z \pm \sqrt{z^2+1})$

21. $\cosh^{-1} z = \ln(z \pm \sqrt{z^2-1}) = \pm\ln(z + \sqrt{z^2-1})$

22. $\tanh^{-1} z = \dfrac{1}{2}\ln\dfrac{1+z}{1-z}$

23. $\cos iz = \cosh z$

24. $\cosh iz = \cos z$

25. (a) Show that $\overline{\cos z} = \cos \bar{z}$.

(b) Is $\overline{\sin z} = \sin \bar{z}$?

(c) If $f(z) = 1 + iz$, is $\overline{f(z)} = f(\bar{z})$?

(d) If $f(z)$ is expanded in a power series with *real* coefficients, show that $\overline{f(z)} = f(\bar{z})$.

(e) Using part (d), verify, *without computing its value*, that $i[\sinh(1+i) - \sinh(1-i)]$ is real.

26. Find $\left|\dfrac{2e^{i\theta}-i}{ie^{i\theta}+2}\right|$. *Hint:* See equation (5.1).

27. (a) Show that $\operatorname{Re} z = \frac{1}{2}(z+\bar{z})$ and that $\operatorname{Im} z = (1/2i)(z-\bar{z})$.

(b) Show that $|e^z|^2 = e^{2\operatorname{Re} z}$.

(c) Use (b) to evaluate $|e^{(1+ix)^2(1-it)-|1+it|^2}|^2$ which occurs in quantum mechanics.

28. Evaluate the following absolute square of a complex number (which arises in a problem in quantum mechanics). Assume a and b are real. Express your answer in terms of a hyperbolic function.

$$\left|\frac{(a+bi)^2e^b - (a-bi)^2e^{-b}}{4abie^{-ia}}\right|^2$$

29. If $z = \dfrac{a}{b}$ and $\dfrac{1}{a+b} = \dfrac{1}{a} + \dfrac{1}{b}$, find z.

30. Write the series for $e^{x(1+i)}$. Write $1+i$ in the $re^{i\theta}$ form and so obtain (easily) the powers of $(1+i)$. Thus show, for example, that the $e^x \cos x$ series has no x^2 term, no x^6 term, etc., and a similar result for the $e^x \sin x$ series. Find (easily) a formula for the general term for each series.

31. Show that if a sequence of complex numbers tends to zero, then the sequence of absolute values tends to zero too, and vice versa. *Hint:* $a_n + ib_n \to 0$ means $a_n \to 0$ and $b_n \to 0$.

32. Use a series you know to show that $\displaystyle\sum_{n=0}^{\infty} \frac{(1+i\pi)^n}{n!} = -e$.

FĚUE WHEP

CHAPTER 3

Linear Algebra

1. INTRODUCTION

In this chapter, we are going to discuss a combination of algebra and geometry which is important in many applications. As you know, problems in various fields of science and mathematics involve the solution of sets of linear equations. This sounds like algebra, but it has a useful geometric interpretation. Suppose you have solved two simultaneous linear equations and have found $x = 2$ and $y = -3$. We can think of $x = 2$, $y = -3$ as the point $(2, -3)$ in the (x, y) plane. Since two linear equations represent two straight lines, the solution is then the point of intersection of the lines. The geometry helps us to understand that sometimes there is no solution (parallel lines) and sometimes there are infinitely many solutions (both equations represent the same line).

The language of vectors is very useful in studying sets of simultaneous equations. You are familiar with quantities such as the velocity of an object, the force acting on it, or the magnetic field at a point, which have both magnitude and direction. Such quantities are called *vectors*; contrast them with such quantities as mass, time, or temperature, which have magnitude only and are called *scalars*. A vector can be represented by an arrow and labeled by a boldface letter ($\mathbf{A}$ in Figure 1.1; also see Section 4). The length of the arrow tells us the magnitude of the vector and the direction of the arrow tells us the direction of the vector. It is not necessary to use coordinate axes as in Figure 1.1; we can, for example, point a finger to tell someone which way it is to town without knowing the direction of north. This is the geometric method of discussing vectors (see Section 4). However, if we do use a coordinate system as in Figure 1.1, we can specify the vector by giving its *components* A_x and A_y which are the projections of the vector on the x axis and the y axis. Thus we have two distinct methods of defining and working with vectors. A vector may be a geometric entity (arrow), or it may be a set of numbers (components relative to a coordinate system) which we use algebraically. As we shall see, this double interpretation of everything we do makes the use of vectors a very powerful tool in applications.

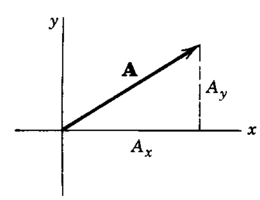


Figure 1.1

One of the great advantages of vector formulas is that they are independent of

10325

the choice of coordinate system. For example, suppose we are discussing the motion of a mass m sliding down an inclined plane. Newton's second law $\mathbf{F} = m\mathbf{a}$ is then a correct equation no matter how we choose our axes. We might, say, take the x axis horizontal and the y axis vertical, or alternatively we might take the x axis along the inclined plane and the y axis perpendicular to the plane. F_x would, of course, be different in the two cases, but for either case it would be true that $F_x = ma_x$ and $F_y = ma_y$, that is, the *vector* equation $\mathbf{F} = m\mathbf{a}$ would be true.

As we have just seen, a vector equation in two dimensions is equivalent to two component equations. In three dimensions, a vector equation is equivalent to three component equations. We will find it useful to generalize this to n dimensions and think of a set of n equations in n unknowns as the component equations for a vector equation in an n dimensional space (Section 10).

We shall also be interested in sets of linear equations which you can think of as changes of variable, say

$$\begin{cases} x' = ax + by, \\ y' = cx + dy, \end{cases} \tag{1.1}$$

where a, b, c, d, are constants. Alternatively, we can think of (1.1) geometrically as telling us to move each point (x, y) to another point (x', y'), an operation we will refer to as a transformation of the plane. Or if we think of (x, y) and (x', y') as being components of vectors from the origin to the given points, then (1.1) tells us how to change each vector in the plane to another vector. Equations (1.1) could also correspond to a change of axes (say a rotation of axes around the origin) where (x, y) and (x', y') are the coordinates of the same point relative to different axes. We will learn (Sections 11 and 12) how to choose the best coordinate system or set of variables to use in solving various problems. The same methods and tools (such as matrices and determinants) which can be used to solve sets of numerical equations are what we need to work with transformations and changes of coordinate system. After we have considered 2- and 3-dimensional space, we will extend these ideas to n-dimensional space and finally to a space in which the "vectors" are functions. This generalization is of great importance in applications.

2. MATRICES; ROW REDUCTION

A matrix (plural: matrices) is just a rectangular array of quantities, usually inclosed in large parentheses, such as

$$\mathrm{A} = \begin{pmatrix} 1 & 5 & -2 \\ -3 & 0 & 6 \end{pmatrix}. \tag{2.1}$$

We will ordinarily indicate a matrix by a roman letter such as A (or B, C, M, r, etc.), but the letter does not have a numerical value; it simply stands for the array. To indicate a number in the array, we will write A_{ij} where i is the row number and j is the column number. For example, in (2.1), $A_{11} = 1$, $A_{12} = 5$, $A_{13} = -2$, $A_{21} = -3$, $A_{22} = 0$, $A_{23} = 6$. We will call a matrix with m rows and n columns an m by n matrix. Thus the matrix in (2.1) is a 2 by 3 matrix, and the matrix in (2.2) below is a 3 by 2 matrix.

FÊUE WHEP
10326

Transpose of a Matrix We write

$$\mathrm{A}^{\mathrm{T}} = \begin{pmatrix} 1 & -3 \\ 5 & 0 \\ -2 & 6 \end{pmatrix}, \tag{2.2}$$

and call A^{T} the transpose of the matrix A in (2.1). To transpose a matrix, we simply write the rows as columns, that is, we interchange rows and columns. Note that, using index notation, we have $(\mathrm{A}^{\mathrm{T}})_{ij} = A_{ji}$. You will find a summary of matrix notation in Section 9.

Sets of Linear Equations Historically linear algebra grew out of efforts to find efficient methods for solving sets of linear equations. As we have said, the subject has developed far beyond the solution of sets of numerical equations (which are easily solved by computer), but the ideas and methods developed for that purpose are needed in later work. A simple way to learn these techniques is to use them to solve some numerical problems by hand. In this section and the next we will develop methods of working with sets of linear equations, and introduce definitions and notation which will be useful later. Also, as you will see, we will discover how to tell whether a given set of equations has a solution or not.

Example 1. Consider the set of equations

$$\begin{cases} 2x \phantom{{}+5y} - z = 2, \\ 6x + 5y + 3z = 7, \\ 2x - y \phantom{{}+3z} = 4. \end{cases} \tag{2.3}$$

Let's agree always to write sets of equations in this *standard form* with the x terms lined up in a column (and similarly for the other variables), and with the constants on the right hand sides of the equations. Then there are several matrices of interest connected with these equations. First is the *matrix of the coefficients* which we will call M:

$$\mathrm{M} = \begin{pmatrix} 2 & 0 & -1 \\ 6 & 5 & 3 \\ 2 & -1 & 0 \end{pmatrix}. \tag{2.4}$$

Then there are two 3 by 1 matrices which we will call r and k:

$$\mathrm{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \mathrm{k} = \begin{pmatrix} 2 \\ 7 \\ 4 \end{pmatrix}. \tag{2.5}$$

If we use index notation and replace x, y, z, by x_1, x_2, x_3, and call the constants k_1, k_2, k_3, then we could write the equations (2.3) in the form (Problem 1)

$$\sum_{j=1}^{3} M_{ij}x_j = k_i, \quad i = 1, 2, 3. \tag{2.6}$$

It is interesting to note that, as we will see in Section 6, this is exactly how matrices are multiplied, so we will learn to write sets of equations like (2.3) as $\mathrm{Mr} = \mathrm{k}$.

FEUE WHEP

For right now we are interested in the fact that we can display all the essential numbers in equations (2.3) as a matrix known as the *augmented matrix* which we call A. Note that the first three columns of A are just the columns of M, and the fourth column is the column of constants on the right hand sides of the equations.

$$\text{A} = \begin{pmatrix} 2 & 0 & -1 & 2 \\ 6 & 5 & 3 & 7 \\ 2 & -1 & 0 & 4 \end{pmatrix}. \tag{2.7}$$

Instead of working with a set of equations and writing all the variables, we can just work with the matrix (2.7). The process which we are going to show is called *row reduction* and is essentially the way your computer solves a set of linear equations. Row reduction is just a systematic way of taking linear combinations of the given equations to produce a simpler but equivalent set of equations. We will show the process, writing side-by-side the equations and the matrix corresponding to them.

(a) The first step is to use the first equation in (2.3) to eliminate the x terms in the other two equations. The corresponding matrix operation on (2.7) is to subtract 3 times the first row from the second row and subtract the first row from the third row. This gives:

$$\begin{cases} 2x \qquad\quad - z = 2, \\ \qquad 5y + 6z = 1, \\ \quad - y + z = 2. \end{cases} \qquad \begin{pmatrix} 2 & 0 & -1 & 2 \\ 0 & 5 & 6 & 1 \\ 0 & -1 & 1 & 2 \end{pmatrix}$$

(b) Now it is convenient to interchange the second and third equations to get:

$$\begin{cases} 2x \qquad\quad - z = 2, \\ \quad - y + z = 2, \\ \qquad 5y + 6z = 1. \end{cases} \qquad \begin{pmatrix} 2 & 0 & -1 & 2 \\ 0 & -1 & 1 & 2 \\ 0 & 5 & 6 & 1 \end{pmatrix}$$

(c) Next we use the second equation to eliminate the y terms from the other equations:

$$\begin{cases} 2x \qquad\quad - z = 2, \\ \quad - y + z = 2, \\ \qquad\qquad 11z = 11. \end{cases} \qquad \begin{pmatrix} 2 & 0 & -1 & 2 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 11 & 11 \end{pmatrix}$$

(d) Finally, we divide the third equation by 11 and then use it to eliminate the z terms from the other equations:

$$\begin{cases} 2x \qquad = 3, \\ \quad - y \quad = 1, \\ \qquad\quad z = 1. \end{cases} \qquad \begin{pmatrix} 2 & 0 & 0 & 3 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

It is customary to divide each equation by the leading coefficient so that the equations read $x = 3/2$, $y = -1$, $z = 1$. The row reduced matrix is then:

$$\begin{pmatrix} 1 & 0 & 0 & 3/2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

The important thing to understand here is that in finding a row reduced matrix we have just taken linear combinations of the original equations. This process is

FÊUE WHEP 10328

reversible, so the final simple equations are equivalent to the original ones. Let's summarize the allowed operations in row reducing a matrix (called *elementary row operations*).

(2.8) i. Interchange two rows [see step (b)];
ii. Multiply (or divide) a row by a (nonzero) constant [see step (d)];
iii. Add a multiple of one row to another; this includes subtracting, that is, using a negative multiple [see steps (a) and (c)].

▶ **Example 2.** Write and row reduce the augmented matrix for the equations:

$$\left\{\begin{aligned} x - y + 4z &= 5,\\ 2x - 3y + 8z &= 4,\\ x - 2y + 4z &= 9. \end{aligned}\right. \tag{2.9}$$

This time we won't write the equations, just the augmented matrix. Remember the routine: Use the first row to clear the rest of the first column; use the new second row to clear the rest of the second column; etc. Also, since matrices are equal only if they are identical, we will not use equal signs between them. Let's use arrows.

$$\begin{pmatrix} 1 & -1 & 4 & 5\\ 2 & -3 & 8 & 4\\ 1 & -2 & 4 & 9 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 4 & 5\\ 0 & -1 & 0 & -6\\ 0 & -1 & 0 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 4 & 11\\ 0 & -1 & 0 & -6\\ 0 & 0 & 0 & -20 \end{pmatrix}$$

We don't need to go any farther! The last row says $0 \cdot z = -20$ which isn't true for any finite value of z. Now you see why your computer doesn't give an answer—there isn't any. We say that the equations are *inconsistent*. If this happens for a set of equations you have written for a physics problem, you know to look for a mistake.

Rank of a Matrix There is another way to discuss Example 2 using the following definition: The number of nonzero rows remaining when a matrix has been row reduced is called the *rank* of the matrix. (It is a theorem that the rank of A^{T} is the same as the rank of A.) Now look at the reduced augmented matrix for Example 2; it has 3 nonzero rows so its rank is 3. But the matrix M (matrix of the coefficients = first three columns of A) has only 2 nonzero rows so its rank is 2. Note that (rank of M) $<$ (rank of A) and the equations are inconsistent.

▶ **Example 3.** Consider the equations

$$\left\{\begin{aligned} x + 2y - z &= 4,\\ 2x \phantom{{}+2y} - z &= 1,\\ x - 2y \phantom{{}-z} &= -3. \end{aligned}\right. \tag{2.10}$$

Either by hand or by computer we row reduce the augmented matrix to get:

$$\begin{pmatrix} 1 & 2 & -1 & 4\\ 2 & 0 & -1 & 1\\ 1 & -2 & 0 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1/2 & 1/2\\ 0 & 1 & -1/4 & 7/4\\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

10329

The last row of zeros tells us that there are infinitely many solutions. For any z we find from the first two rows that $x = (z+1)/2$ and $y = (z+7)/4$. Here we see that the rank of M and the rank of A are both 2 but the number of unknowns is 3, and we are able to find two unknowns in terms of the third.

To make this all very clear, let's look at some simple examples where the results are obvious. We write three sets of equations together with the row reduced matrices:

$$(2.11) \qquad \begin{cases} x + y = 2, \\ x + y = 5. \end{cases} \qquad \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

$$(2.12) \qquad \begin{cases} x + y = 2, \\ 2x + 2y = 4. \end{cases} \qquad \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(2.13) \qquad \begin{cases} x + y = 2, \\ x - y = 4. \end{cases} \qquad \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \end{pmatrix}$$

In (2.11), since $x + y$ can't be equal to both 2 and 5, it is clear that there is no solution; the equations are inconsistent. Note that the last row of the reduced matrix is all zeros except for the last entry and so (rank M) $<$ (rank A). In (2.12), the second equation is just twice the first so they are really the same equation; we say that the equations are dependent. There is an infinite set of solutions, namely all points on the line $y = 2-x$. Note that the last line of the matrix is all zeros; this indicates linear dependence. We have (rank A) $=$ (rank M) $= 1$, and we can solve for one unknown in terms of the other. Finally in (2.13) we have a set of equations with one solution, $x = 3$, $y = -1$, and we see that the row reduced matrix gives this result. Note that (rank A) $=$ (rank M) $=$ number of unknowns $= 2$.

Now let's consider the general problem of solving m equations in n unknowns. Then M has m rows (corresponding to m equations) and n columns (corresponding to n unknowns) and A has one more column (the constants). The following summary outlines the possible cases.

(2.14)

a. If (rank M) $<$ (rank A), the equations are inconsistent and there is no solution.

b. If (rank M) $=$ (rank A) $= n$ (number of unknowns), there is one solution.

c. If (rank M) $=$ (rank A) $= R < n$, then R unknowns can be found in terms of the remaining $n - R$ unknowns.

Example 4. Here is a set of equations and the row reduced matrix:

$$(2.15) \qquad \begin{cases} x + y - z = 7, \\ 2x - y - 5z = 2, \\ -5x + 4y + 14z = 1, \\ 3x - y - 7z = 5. \end{cases} \qquad \begin{pmatrix} 1 & 0 & -2 & 3 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

10330

From the reduced matrix, the solution is $x = 3 + 2z$, $y = 4 - z$. We see that this is an example of (2.14c) with $m = 4$ (number of equations), $n = 3$ (number of unknowns), (rank M) = (rank A) = $R = 2 < n = 3$. Then by (2.14c), we solve for $R = 2$ unknowns (x and y) in terms of the $n - R = 1$ unknown (z).

PROBLEMS, SECTION 2

1. The first equation in (2.6) written out in detail is

$$M_{11}x_1 + M_{12}x_2 + M_{13}x_3 = k_1.$$

Write out the other two equations in the same way and then substitute $x_1, x_2, x_3 = x, y, z$ and the values of M_{ij} and k_i from (2.4) and (2.5) to verify that (2.6) is really (2.3).

2. As in Problem 1, write out in detail in terms of M_{ij}, x_j, and k_i, equations like (2.6) for two equations in four unknowns; for four equations in two unknowns.

For each of the following problems write and row reduce the augmented matrix to find out whether the given set of equations has exactly one solution, no solutions, or an infinite set of solutions. Check your results by computer. *Warning hint:* Be sure your equations are written in standard form. *Comment*: Remember that the point of doing these problems is not just to get an answer (which your computer will give you), but to become familiar with the terminology, ideas, and notation we are using.

3. $\begin{cases} x - 2y + 13 = 0 \\ y - 4x = 17 \end{cases}$

4. $\begin{cases} 2x + y - z = 2 \\ 4x + y - 2z = 3 \end{cases}$

5. $\begin{cases} 2x + y - z = 2 \\ 4x + 2y - 2z = 3 \end{cases}$

6. $\begin{cases} x + y - z = 1 \\ 3x + 2y - 2z = 3 \end{cases}$

7. $\begin{cases} 2x + 3y = 1 \\ x + 2y = 2 \\ x + 3y = 5 \end{cases}$

8. $\begin{cases} -x + y - z = 4 \\ x - y + 2z = 3 \\ 2x - 2y + 4z = 6 \end{cases}$

9. $\begin{cases} x - y + 2z = 5 \\ 2x + 3y - z = 4 \\ 2x - 2y + 4z = 6 \end{cases}$

10. $\begin{cases} x + 2y - z = 1 \\ 2x + 3y - 2z = -1 \\ 3x + 4y - 3z = -4 \end{cases}$

11. $\begin{cases} x - 2y = 4 \\ 5x + z = 7 \\ x + 2y - z = 3 \end{cases}$

12. $\begin{cases} 2x + 5y + z = 2 \\ x + y + 2z = 1 \\ x + 5z = 3 \end{cases}$

13. $\begin{cases} 4x + 6y - 12z = 7 \\ 5x - 2y + 4z = -15 \\ 3x + 4y - 8z = 4 \end{cases}$

14. $\begin{cases} 2x + 3y - z = -2 \\ x + 2y - z = 4 \\ 4x + 7y - 3z = 11 \end{cases}$

Find the rank of each of the following matrices.

15. $\begin{pmatrix} 1 & 1 & 2 \\ 2 & 4 & 6 \\ 3 & 2 & 5 \end{pmatrix}$

16. $\begin{pmatrix} 2 & -3 & 5 & 3 \\ 4 & -1 & 1 & 1 \\ 3 & -2 & 3 & 4 \end{pmatrix}$

17. $\begin{pmatrix} 1 & 1 & 4 & 3 \\ 3 & 1 & 10 & 7 \\ 4 & 2 & 14 & 10 \\ 2 & 0 & 6 & 4 \end{pmatrix}$

18. $\begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & -2 & -1 & 0 \\ 2 & 2 & 5 & 3 \\ 2 & 4 & 8 & 6 \end{pmatrix}$

FEUE WHEP
10331

3. DETERMINANTS; CRAMER'S RULE

We have said that a matrix is simply a display of a set of numbers; it does *not* have a numerical value. For a square matrix, however, there is a useful number called the *determinant* of the matrix. Although a computer will quickly give the value of a determinant, we need to know what this value means in order to use it in applications. [See, for example, equations (4.19), (6.24) and (8.5).] We also need to know how to work with determinants. An easy way to learn these things is to solve some numerical problems by hand. We shall outline some of the facts about determinants without proofs (for more details, see linear algebra texts).

Evaluating Determinants To indicate that we mean the determinant of a square matrix A (written det A), we replace the large parentheses inclosing A by single bars. The value of det A if A is a 1 by 1 matrix is just the value of the single element. For a 2 by 2 matrix,

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \det \mathrm{A} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{3.1}$$

Equation (3.1) gives the value of a second order determinant. We shall describe how to evaluate determinants of higher order.

First we need some notation and definitions. It is convenient to write an n^{th} order determinant like this:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & & & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}. \tag{3.2}$$

Notice that a_{23} is the element in the second row and the third column; that is, the first subscript is the number of the row and the second subscript is the number of the column in which the element is. Thus the element a_{ij} is in row i and column j. As an abbreviation for the determinant in (3.2), we sometimes write simply $|a_{ij}|$, that is, the determinant whose elements are a_{ij}. In this form it looks exactly like the absolute value of the element a_{ij} and you have to tell from the context which of these meanings is intended.

If we remove one row and one column from a determinant of order n, we have a determinant of order $n-1$. Let us remove the row and column containing the element a_{ij} and call the remaining determinant M_{ij}. The determinant M_{ij} is called the *minor* of a_{ij}. For example, in the determinant

$$\begin{vmatrix} 1 & -5 & 2 \\ 7 & 3 & 4 \\ 2 & 1 & 5 \end{vmatrix}, \tag{3.3}$$

the minor of the element $a_{23} = 4$ is

$$M_{23} = \begin{vmatrix} 1 & -5 \\ 2 & 1 \end{vmatrix},$$

FEUE WHEP
10332

obtained by crossing off the row and column containing 4. The signed minor $(-1)^{i+j}M_{ij}$ is called the *cofactor* of a_{ij}. In (3.3), the element 4 is in the second row ($i = 2$) and third column ($j = 3$), so $i + j = 5$, and the cofactor of 4 is $(-1)^5 M_{23} = -11$. It is very convenient to get the proper sign (plus or minus) for the factor $(-1)^{i+j}$ by thinking of a checkerboard of plus and minus signs like this:

$$\tag{3.4} \begin{vmatrix} + & - & + & - & & & \\ - & + & - & + & & & \\ + & - & + & - & \text{etc.} & & \\ - & + & - & + & & & \\ & \text{etc.} & & & \ddots & & \\ & & & & & + & - \\ & & & & & - & + \end{vmatrix}.$$

Then the sign $(-1)^{i+j}$ to be attached to M_{ij} is just the checkerboard sign in the same position as a_{ij}. For the element a_{23}, you can see that the checkerboard sign is minus.

Now we can easily say how to find the *value of a determinant: Multiply each element of one row (or one column) by its cofactor and add the results.* It can be shown that we get the same answer whichever row or column we use.

▶ **Example 1.** Let us evaluate the determinant in (3.3) using elements of the third column. We get

$$\begin{vmatrix} 1 & -5 & 2 \\ 7 & 3 & 4 \\ 2 & 1 & 5 \end{vmatrix} = 2\begin{vmatrix} 7 & 3 \\ 2 & 1 \end{vmatrix} - 4\begin{vmatrix} 1 & -5 \\ 2 & 1 \end{vmatrix} + 5\begin{vmatrix} 1 & -5 \\ 7 & 3 \end{vmatrix}$$
$$= 2 \cdot 1 - 4 \cdot 11 + 5 \cdot 38 = 148.$$

As a check, using elements of the first row, we get

$$1\begin{vmatrix} 3 & 4 \\ 1 & 5 \end{vmatrix} + 5\begin{vmatrix} 7 & 4 \\ 2 & 5 \end{vmatrix} + 2\begin{vmatrix} 7 & 3 \\ 2 & 1 \end{vmatrix} = 11 + 135 + 2 = 148.$$

The method of evaluating a determinant which we have described here is one form of Laplace's development of a determinant. If the determinant is of fourth order (or higher), using the Laplace development once gives us a set of determinants of order one less than we started with; then we use the Laplace development all over again to evaluate each of these, and so on until we get determinants of second order which we know how to evaluate. This is obviously a lot of work! We will see below how to simplify the calculation. A word of warning to anyone who has learned a special method of evaluating a third-order determinant by recopying columns to the right and multiplying along diagonals: this method *does not work* for fourth order (and higher).

FEUE WHEP

Useful Facts About Determinants We state these facts without proof. (See algebra books for proofs.)

1. If each element of *one* row (or *one* column) of a determinant is multiplied by a number k, the value of the determinant is multiplied by k.
2. The value of a determinant is zero if
   (a) all elements of one row (or column) are zero; or if
   (b) two rows (or two columns) are identical; or if
   (c) two rows (or two columns) are proportional.
3. If two rows (or two columns) of a determinant are interchanged, the value of the determinant changes sign.
4. The value of a determinant is unchanged if
   (a) rows are written as columns and columns as rows; or if
   (b) we add to each element of one row, k times the corresponding element of another row, where k is any number (and a similar statement for columns).

Let us look at a few examples of the use of these facts.

Example 2. Find the equation of a plane through the three given points $(0,0,0)$, $(1,2,5)$, and $(2,-1,0)$.

We shall verify that the answer in determinant form is

$$\begin{vmatrix} x & y & z & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 5 & 1 \\ 2 & -1 & 0 & 1 \end{vmatrix} = 0.$$

By a Laplace development using elements of the first row, we would find that this is a linear equation in x, y, z; thus it represents a plane. We need now to show that the three points are in the plane. Suppose $(x,y,z) = (0,0,0)$; then the first two rows of the determinant are identical and by Fact 2b the determinant is zero. Similarly if the point (x,y,z) is either of the other given points, two rows of the determinant are identical and the determinant is zero. Thus all three points lie in the plane.

Example 3. Evaluate the determinant

$$D = \begin{vmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{vmatrix}.$$

10334

If we interchange rows and columns in D, then by Facts 4a and 1 we have

$$D = \begin{vmatrix} 0 & -a & b \\ a & 0 & -c \\ -b & c & 0 \end{vmatrix} = (-1)^3 \begin{vmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{vmatrix},$$

where in the last step we have factored -1 out of each column by Fact 1. Thus we have $D = -D$, so $D = 0$.

We can use Facts 1 to 4 to simplify finding the value of a determinant. First we check Facts 2a, 2b, 2c, in case the determinant is trivially equal to zero. Then we try to get as many zeros as possible in some row or column in order to have fewer terms in the Laplace development. We look for rows (or columns) which can be combined (using Fact 4b) to give zeros. Although this is something like row reduction, we can operate with columns as well as rows. However, we can't just cancel a number from a row (or column); by Fact 1 we must keep it as a factor in our answer. And we must keep track of any row (or column) interchanges since by Fact 3 each interchange multiplies the determinant by (-1).

▶ **Example 4.** Evaluate the determinant

$$D = \begin{vmatrix} 4 & 3 & 0 & 1 \\ 9 & 7 & 2 & 3 \\ 4 & 0 & 2 & 1 \\ 3 & -1 & 4 & 0 \end{vmatrix}.$$

Subtract 4 times the fourth column from the first column, and subtract 2 times the fourth column from the third column to get:

$$D = \begin{vmatrix} 0 & 3 & -2 & 1 \\ -3 & 7 & -4 & 3 \\ 0 & 0 & 0 & 1 \\ 3 & -1 & 4 & 0 \end{vmatrix}.$$

Do a Laplace development using the third row:

$$D = (-1) \begin{vmatrix} 0 & 3 & -2 \\ -3 & 7 & -4 \\ 3 & -1 & 4 \end{vmatrix}. \tag{3.5}$$

Add the second row to the third row:

$$D = (-1) \begin{vmatrix} 0 & 3 & -2 \\ -3 & 7 & -4 \\ 0 & 6 & 0 \end{vmatrix}.$$

Do a Laplace development using the first column:

$$D = (-1)(-1)(-3) \begin{vmatrix} 3 & -2 \\ 6 & 0 \end{vmatrix} = (-3)[0 - 6(-2)] = -36.$$

This is the answer but you might like to look for some shorter solutions. For example, consider the determinant (3.5) above. If we immediately do another Laplace development using the first row, the minor of 3 in the first row, second column is

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

10335

Without even evaluating it, we should recognize by Fact 2c that it is zero. Then proceeding with the Laplace development of (3.5) using the first row gives just

$$D = (-1)(-2)\begin{vmatrix} -3 & 7 \\ 3 & -1 \end{vmatrix} = 2(3-21) = -36 \quad \text{as above.}$$

Now you may be wondering why you should learn about this when your computer will do it for you. Suppose you have a determinant with elements which are algebraic expressions, and you want to write it in a different form. Then you need to know what manipulations you can do without changing its value. Also, if you know the rules, you may see that a determinant is zero without evaluating it. An easy way to learn these things is to evaluate some simple numerical determinants by hand.

Cramer's Rule This is a formula in terms of determinants for the solution of n linear equations in n unknowns when there is exactly one solution. As we said for row reduction and for evaluating determinants, your computer will quickly give you the solution of a set of linear equations when there is one. However, for theoretical purposes, we need the Cramer's rule formula, and a simple way to learn about it is to use it to solve sets of linear equations with numerical coefficients.

Let us first show the use of Cramer's rule to solve two equations in two unknowns. Then we will generalize it to n equations in n unknowns. Consider the set of equations

$$\begin{cases} a_1x + b_1y = c_1, \\ a_2x + b_2y = c_2. \end{cases} \tag{3.6}$$

If we multiply the first equation by b_2, the second by b_1, and then subtract the results and solve for x, we get (if $a_1b_2 - a_2b_1 \neq 0$)

$$x = \frac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1}. \tag{3.7a}$$

Solving for y in a similar way, we get

$$y = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}. \tag{3.7b}$$

Using the definition (3.1) of a second order determinant, we can write the solutions (3.7) of (3.6) in the form

$$x = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \qquad y = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}. \tag{3.8}$$

It is helpful in remembering (3.8) to say in words how we find the correct determinants. First, the equations must be written in standard form as for row reduction (Section 2). Then if we simply write the array of coefficients on the left-hand side of (3.6), these form the denominator determinant in (3.8). This determinant (which we shall denote by D) is called the *determinant of the coefficients*. To find the numerator determinant for x, start with D, erase the x coefficients a_1 and a_2, and replace them by the constants c_1 and c_2 from the right-hand sides of the equations. Similarly, we replace the y coefficients in D by the constant terms to find the numerator determinant in y.

10336

▶ **Example 5.** Use (3.8) to solve the set of equations

$$\begin{cases} 2x + 3y = 3, \\ x - 2y = 5. \end{cases}$$

We find

$$D = \begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -4 - 3 = -7,$$

$$x = \frac{1}{D}\begin{vmatrix} 3 & 3 \\ 5 & -2 \end{vmatrix} = \frac{-6-15}{-7} = 3,$$

$$y = \frac{1}{D}\begin{vmatrix} 2 & 3 \\ 1 & 5 \end{vmatrix} = \frac{10-3}{-7} = -1.$$

This method of solution of a set of linear equations is called Cramer's rule. It may be used to solve n equations in n unknowns if $D \neq 0$; the solution then consists of one value for each unknown. The denominator determinant D is the n by n determinant of the coefficients when the equations are arranged in standard form. The numerator determinant for each unknown is the determinant obtained by replacing the column of coefficients of that unknown in D by the constant terms from the right-hand sides of the equations. Then to find the unknowns, we must evaluate each of the determinants and divide.

Rank of a Matrix Here is another way to find the rank of a matrix (Section 2). A submatrix means a matrix remaining if we remove some rows and/or remove some columns from the original matrix. To find the rank of a matrix, we look at all the square submatrices and find their determinants. The order of the largest nonzero determinant is the rank of the matrix.

▶ **Example 6.** Find the rank of the matrix

$$\begin{pmatrix} 1 & -1 & 2 & 3 \\ -2 & 2 & -1 & 0 \\ 4 & -4 & 5 & 6 \end{pmatrix}.$$

We need to look at the four 3 by 3 determinants containing columns 1,2,3 or 1,2,4 or 1,3,4 or 2,3,4. We note that the first two columns are negatives of each other, so by Fact 2c the first two of these determinants are both zero. The last two determinants differ only in the sign of their first column, so we just have to look at one of them, say:

$$\begin{pmatrix} 1 & 2 & 3 \\ -2 & -1 & 0 \\ 4 & 5 & 6 \end{pmatrix}.$$

If we now subtract twice the first row from the third row, we have

$$\begin{pmatrix} 1 & 2 & 3 \\ -2 & -1 & 0 \\ 2 & 1 & 0 \end{pmatrix},$$

10337

and we see by Fact 2c that the determinant is zero. So the rank of the matrix is less than 3. To show that it is 2, we just have to find *one* 2 by 2 submatrix with nonzero determinant. There are several of them; find one. Thus the rank of the matrix is 2. (If we had needed to show that the rank was 1, we would have had to show that *all* the 2 by 2 submatrices had determinants equal to zero.)

PROBLEMS, SECTION 3

Evaluate the determinants in Problems 1 to 6 by the methods shown in Example 4. Remember that the reason for doing this is not just to get the answer (your computer can give you that) but to learn how to manipulate determinants correctly. Check your answers by computer.

1. $\begin{vmatrix} -2 & 3 & 4 \\ 3 & 4 & -2 \\ 5 & 6 & -3 \end{vmatrix}$ **2.** $\begin{vmatrix} 5 & 17 & 3 \\ 2 & 4 & -3 \\ 11 & 0 & 2 \end{vmatrix}$

3. $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix}$ **4.** $\begin{vmatrix} -2 & 4 & 7 & 3 \\ 8 & 2 & -9 & 5 \\ -4 & 6 & 8 & 4 \\ 2 & -9 & 3 & 8 \end{vmatrix}$

5. $\begin{vmatrix} 7 & 0 & 1 & -3 & 5 \\ 2 & -1 & 0 & 1 & 4 \\ 7 & -3 & 2 & -1 & 4 \\ 8 & 6 & -2 & -7 & 4 \\ 1 & 3 & -5 & 7 & 5 \end{vmatrix}$ **6.** $\begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix}$

7. Prove the following by appropriate manipulations using Facts 1 to 4; do not just evaluate the determinants.

$$\begin{vmatrix} 1 & a & bc \\ 1 & b & ac \\ 1 & c & ab \end{vmatrix} = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (c-a)(b-a)(c-b)\begin{vmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 0 & 1 \end{vmatrix}$$
$$= (c-a)(b-a)(c-b).$$

8. Show that if, in using the Laplace development, you accidentally multiply the elements of one row by the cofactors of another row, you get zero.
Hint: Consider Fact 2b.

9. Show without computation that the following determinant is equal to zero.
Hint: Consider the effect of interchanging rows and columns.

$$\begin{vmatrix} 0 & 2 & -3 \\ -2 & 0 & 4 \\ 3 & -4 & 0 \end{vmatrix}$$

10. A determinant or a square matrix is called skew-symmetric if $a_{ij} = -a_{ji}$. (The determinant in Problem 9 is an example of a skew-symmetric determinant.) Show that a skew-symmetric determinant of odd order is zero.

In Problems 11 and 12 evaluate the determinants.

11. $\begin{vmatrix} 0 & 5 & -3 & -4 & 1 \\ -5 & 0 & 2 & 6 & -2 \\ 3 & -2 & 0 & -3 & 7 \\ 4 & -6 & 3 & 0 & -3 \\ -1 & 2 & -7 & 3 & 0 \end{vmatrix}$ **12.** $\begin{vmatrix} 0 & 1 & 2 & -1 \\ -1 & 0 & -3 & 0 \\ -2 & 3 & 0 & 1 \\ 1 & 0 & -1 & 0 \end{vmatrix}$

FEUE WHEP
10338

13. Show that

$$\begin{vmatrix} \cos\theta & 1 & 0 \\ 1 & 2\cos\theta & 1 \\ 0 & 1 & 2\cos\theta \end{vmatrix} = \cos 3\theta.$$

14. Show that the n-rowed determinant

$$\begin{vmatrix} \cos\theta & 1 & 0 & 0 & & & 0 \\ 1 & 2\cos\theta & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 2\cos\theta & 1 & & & 0 \\ 0 & 0 & 1 & 2\cos\theta & & & 0 \\ & & \vdots & & \ddots & & \vdots \\ & & \vdots & & & 2\cos\theta & 1 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2\cos\theta \end{vmatrix} = \cos n\theta.$$

Hint: Expand using elements of the last row or column. Use mathematical induction and the trigonometric addition formulas.

15. Use Cramer's rule to solve Problems 2.3 and 2.11.

16. In the following set of equations (from a quantum mechanics problem), A and B are the unknowns, k and K are given, and $i = \sqrt{-1}$. Use Cramer's rule to find A and show that $|A|^2 = 1$.

$$\begin{cases} A - B = -1 \\ ikA - KB = ik \end{cases}$$

17. Use Cramer's rule to solve for x and t the Lorentz equations of special relativity:

$$\begin{cases} x' = \gamma(x - vt) \\ t' = \gamma(t - vx/c^2) \end{cases} \qquad \text{where} \quad \gamma^2(1 - v^2/c^2) = 1$$

Caution: Arrange the equations in standard form.

18. Find z by Cramer's rule:

$$\begin{cases} (a-b)x - (a-b)y + 3b^2 z = 3ab \\ (a+2b)x - (a+2b)y - (3ab+3b^2)z = 3b^2 \\ bx + ay - (2b^2+a^2)z = 0 \end{cases}$$

▶ 4. VECTORS

Notation We shall indicate a vector by a boldface letter (for example, $\mathbf{A}$) and a component of a vector by a subscript (for example A_x is the x component of $\mathbf{A}$), as in Figure 4.1. Since it is not easy to handwrite boldface letters, you should write a vector with an arrow over it (for example, $\vec{A}$). It is very important to indicate clearly whether a letter represents a vector, since, as we shall see below, the same letter in italics (not boldface) is often used with a different meaning.

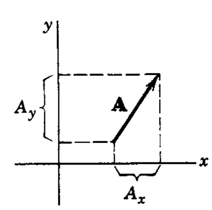


Figure 4.1

Magnitude of a Vector The length of the arrow representing a vector $\mathbf{A}$ is called the *length* or the *magnitude* of $\mathbf{A}$ (written $|\mathbf{A}|$ or A) or (see Section 10) the *norm* of $\mathbf{A}$ (written $||\mathbf{A}||$). Note the use of A to mean the magnitude of $\mathbf{A}$; for this reason it is important to make it clear whether you mean a vector or its magnitude (which is a scalar). By the Pythagorean theorem, we find

10339

$$A = |\mathbf{A}| = \sqrt{A_x^2 + A_y^2} \quad \text{in two dimensions,} \quad \text{or} \tag{4.1}$$
$$A = |\mathbf{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2} \quad \text{in three dimensions.}$$

▶ **Example 1.** In Figure 4.2 the force $\mathbf{F}$ has an x component of 4 lb and a y component of 3 lb. Then we write

$$F_x = 4 \text{ lb},$$
$$F_y = 3 \text{ lb},$$
$$|\mathbf{F}| = 5 \text{ lb},$$
$$\theta = \arctan \tfrac{3}{4}.$$

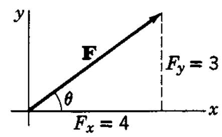


Figure 4.2

Addition of Vectors There are two ways to get the sum of two vectors. One is by the parallelogram law: To find $\mathbf{A} + \mathbf{B}$, place the tail of $\mathbf{B}$ at the head of $\mathbf{A}$ and draw the vector from the tail of $\mathbf{A}$ to the head of $\mathbf{B}$ as shown in Figures 4.3 and 4.4.

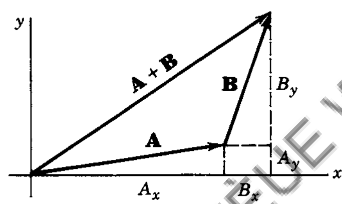


Figure 4.3

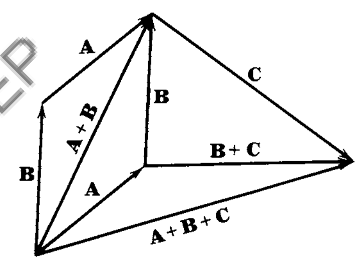


Figure 4.4

The second way of finding $\mathbf{A} + \mathbf{B}$ is to add components: $\mathbf{A} + \mathbf{B}$ has components $A_x + B_x$ and $A_y + B_y$. You should satisfy yourself from Figure 4.3 that these two methods of finding $\mathbf{A} + \mathbf{B}$ are equivalent. From Figure 4.4 and either definition of vector addition, it follows that

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \qquad \text{(commutative law for addition);}$$
$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \qquad \text{(associative law for addition).}$$

In other words, vectors may be added together by the usual laws of algebra.

It seems reasonable to use the symbol $3\mathbf{A}$ for the vector $\mathbf{A} + \mathbf{A} + \mathbf{A}$. By the methods of vector addition above, we can say that the vector $\mathbf{A} + \mathbf{A} + \mathbf{A}$ is a vector three times as long as $\mathbf{A}$ and in the same direction as $\mathbf{A}$ and that each component of $3\mathbf{A}$ is three times the corresponding component of $\mathbf{A}$. As a natural extension of these facts we define the vector $c\mathbf{A}$ (where c is any real positive number) to be a vector c times as long as $\mathbf{A}$ and in the same direction as $\mathbf{A}$; each component of $c\mathbf{A}$ is then c times the corresponding component of $\mathbf{A}$ (Figure 4.5).

The negative of a vector is defined as a vector of the same magnitude but in the opposite direction. Then (Figure 4.6) each component of $-\mathbf{B}$ is the negative of the corresponding component of $\mathbf{B}$. We can now define subtraction of vectors by

10340

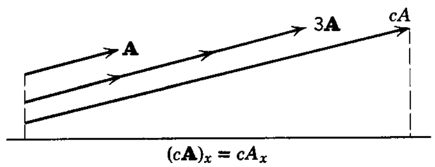


Figure 4.5

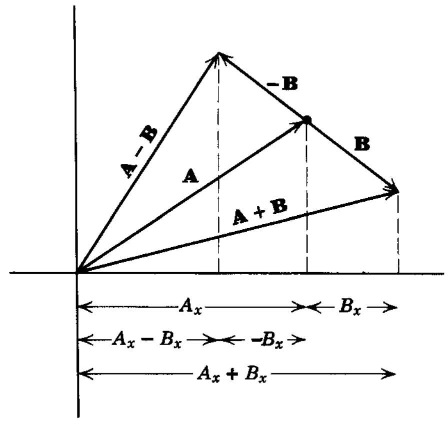


Figure 4.6

saying that $\mathbf{A}-\mathbf{B}$ means the sum of the vectors $\mathbf{A}$ and $-\mathbf{B}$. Each component of $\mathbf{A}-\mathbf{B}$ is then obtained by subtracting the corresponding components of $\mathbf{A}$ and $\mathbf{B}$, that is, $(\mathbf{A}-\mathbf{B})_x = A_x - B_x$, etc. Like addition, subtraction of vectors can be done geometrically (by the parallelogram law) or algebraically by subtracting the components (Figure 4.6).

The *zero vector* (which might arise as $\mathbf{A} = \mathbf{B}-\mathbf{B} = \mathbf{0}$, or as $\mathbf{A} = c\mathbf{B}$ with $c = 0$) is a vector of zero magnitude; its components are all zero and it does not have a direction. A vector of length or magnitude 1 is called a *unit vector*. Then for any $\mathbf{A} \neq 0$, the vector $\mathbf{A}/|\mathbf{A}|$ is a unit vector. In Example 1, $\mathbf{F}/5$ is a unit vector.

We have just seen that there are two ways to combine vectors: geometric (head to tail addition), and algebraic (using components). Let us look first at an example of the geometric method; then we shall consider the algebraic method. Example 2 below illustrates the geometric method. By similar proofs, many of the facts of elementary geometry can be easily proved using vectors, with no reference to components or a coordinate system. (See Problems 3 to 8.)

▶ **Example 2.** Prove that the medians of a triangle intersect at a point two-thirds of the way from any vertex to the midpoint of the opposite side.

To prove this, we call two of the sides of the triangle $\mathbf{A}$ and $\mathbf{B}$. The third side of the triangle is then $\mathbf{A}+\mathbf{B}$ by the parallelogram law, with the directions of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{A}+\mathbf{B}$ as indicated in Figure 4.7. If we add the vector $\frac{1}{2}\mathbf{B}$ to the vector $\mathbf{A}$ (head to tail as in Figure 4.7b), we have a vector from point O to the midpoint of the opposite side of the triangle, that is, we have the median to side $\mathbf{B}$. Next, take two-thirds of this vector; we now have the vector $\frac{2}{3}(\mathbf{A}+\frac{1}{2}\mathbf{B}) = \frac{2}{3}\mathbf{A}+\frac{1}{3}\mathbf{B}$ extending from O to P in Figure 4.7b. We want to show that P is the intersection point of the three medians and also the "$\frac{2}{3}$ point" for each. We prove this by showing that P is the "$\frac{2}{3}$ point" on the median to side $\mathbf{A}$; then since $\mathbf{A}$ and $\mathbf{B}$ represent *any* two sides of the triangle, the proof holds for all three medians. The vector from R to Q (Figure 4.7c) is $\frac{1}{2}\mathbf{A}+\mathbf{B}$; this is the median to $\mathbf{A}$. The "$\frac{2}{3}$ point" on this median is the point P' (Figure 4.7d); the vector from R to P' is equal to $\frac{1}{3}(\frac{1}{2}\mathbf{A}+\mathbf{B})$. Then the vector from O to P' is $\frac{1}{2}\mathbf{A}+\frac{1}{3}(\frac{1}{2}\mathbf{A}+\mathbf{B}) = \frac{2}{3}\mathbf{A}+\frac{1}{3}\mathbf{B}$. Thus P and P' are the same point and all three medians have their "$\frac{2}{3}$ points" there. Note that we have made no reference to a coordinate system or to components in this proof.

FEUE WHEP

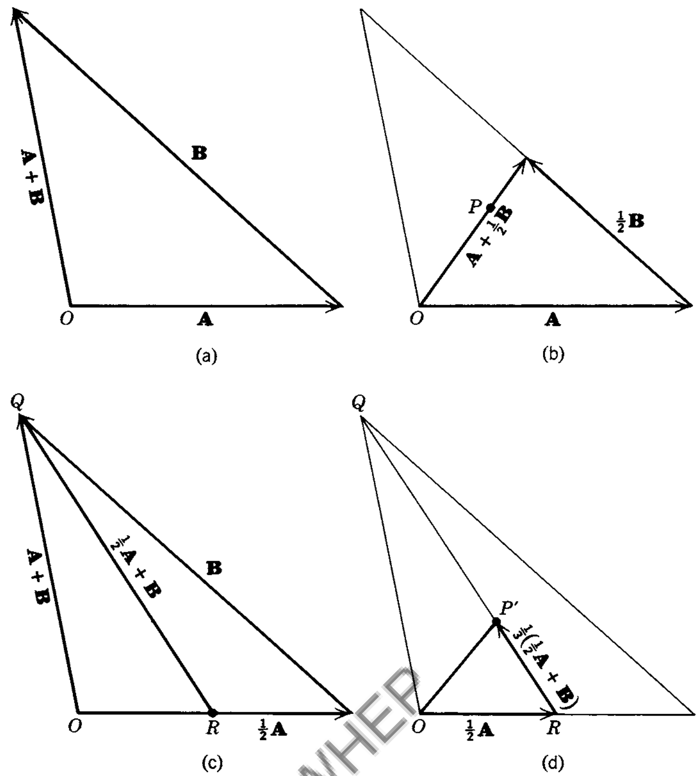


Figure 4.7

PROBLEMS, SECTION 4

1. Draw diagrams and prove (4.1).

2. Given the vectors making the given angles θ with the positive x axis:

$\mathbf{A}$ of magnitude 5, $\theta = 45°$,

$\mathbf{B}$ of magnitude 3, $\theta = -30°$,

$\mathbf{C}$ of magnitude 7, $\theta = 120°$,

(a) Draw diagrams representing $2\mathbf{A}$, $\mathbf{A} - 2\mathbf{B}$, $\mathbf{C} - \mathbf{B}$, $\frac{2}{5}\mathbf{A} - \frac{1}{7}\mathbf{C}$.

(b) Draw diagrams to show that

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \qquad \mathbf{A} - (\mathbf{B} - \mathbf{C}) = (\mathbf{A} - \mathbf{B}) + \mathbf{C},$$
$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = (\mathbf{A} + \mathbf{C}) + \mathbf{B}, \qquad (\mathbf{A} + \mathbf{B})_x = \mathbf{A}_x + \mathbf{B}_x,$$
$$(\mathbf{B} - \mathbf{C})_x = \mathbf{B}_x - \mathbf{C}_x.$$

Use vectors to prove the following theorems from geometry:

3. The diagonals of a parallelogram bisect each other.

4. The line segment joining the midpoints of two sides of any triangle is parallel to the third side and half its length.

5. In a parallelogram, the two lines from one corner to the midpoints of the two opposite sides trisect the diagonal they cross.

10342

6. In any quadrilateral (four-sided figure with sides of various lengths and—in general—four different angles), the lines joining the midpoints of opposite sides bisect each other. *Hint:* Label three sides **A**, **B**, **C**; what is the vector along the fourth side?

7. A line through the midpoint of one side of a triangle and parallel to a second side bisects the third side. *Hint:* Call parallel vectors **A** and $c\mathbf{A}$.

8. The median of a trapezoid (four-sided figure with just two parallel sides) means the line joining the midpoints of the two nonparallel sides. Prove that the median bisects both diagonals; that the median is parallel to the two parallel bases and equal to half the sum of their lengths.

We have discussed in some detail the geometric method of adding vectors (parallelogram law or head to tail addition) and its importance in stating and proving geometric and physical facts without the intrusion of a special coordinate system. There are, however, many cases in which algebraic methods (using components relative to a particular coordinate system) are better. We shall discuss this next.

Vectors in Terms of Components We consider a set of rectangular axes as in Figure 4.8. Let the vector **i** be a unit vector in the positive x direction (out of the paper toward you), and let **j** and **k** be unit vectors in the positive y and z directions. If A_x and A_y are the scalar components of a vector in the (x, y) plane, then $\mathbf{i}A_x$ and $\mathbf{j}A_y$ are its vector components, and their sum is the vector **A** (Figure 4.9).

$$\mathbf{A} = \mathbf{i}A_x + \mathbf{j}A_y.$$

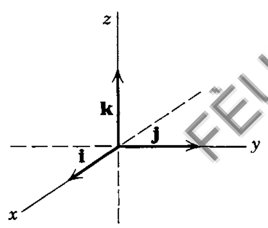


Figure 4.8

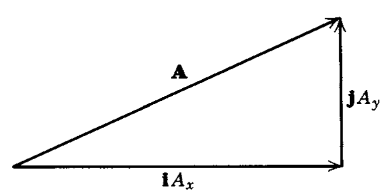


Figure 4.9

Similarly, in three dimensions

$$\mathbf{A} = \mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z.$$

It is easy to add (or subtract) vectors in this form: If **A** and **B** are vectors in two dimensions, then

$$\mathbf{A} + \mathbf{B} = (\mathbf{i}A_x + \mathbf{j}A_y) + (\mathbf{i}B_x + \mathbf{j}B_y) = \mathbf{i}(A_x + B_x) + \mathbf{j}(A_y + B_y).$$

This is just the familiar result of adding components; the unit vectors **i** and **j** serve to keep track of the separate components and allow us to write **A** as a single algebraic expression. The vectors **i**, **j**, **k** are called *unit basis vectors.*

Multiplication of Vectors There are two kinds of product of two vectors. One, called the *scalar product* (or *dot product* or *inner product*), gives a result which is a scalar; the other, called the *vector product* (or *cross product*), gives a vector answer.

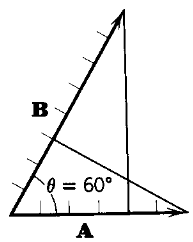
Scalar Product By definition, the scalar product of $\mathbf{A}$ and $\mathbf{B}$ (written $\mathbf{A} \cdot \mathbf{B}$) is a scalar equal to the magnitude of $\mathbf{A}$ times the magnitude of $\mathbf{B}$ times the cosine of the angle θ between $\mathbf{A}$ and $\mathbf{B}$:

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}|\,|\mathbf{B}| \cos\theta. \tag{4.2}$$

You should observe from (4.2) that the commutative law (4.3) holds for scalar multiplication:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}. \tag{4.3}$$

A useful interpretation of the dot product is shown in Figure 4.10.



$|\mathbf{B}| = 8,\ |\mathbf{A}| = 6.$
Projection of $\mathbf{B}$ on $\mathbf{A} = 4$;
$\mathbf{A} \cdot \mathbf{B} = 6 \cdot 4 = 24.$
Or, projection of $\mathbf{A}$ on $\mathbf{B} = 3$;
$\mathbf{B} \cdot \mathbf{A} = 3 \cdot 8 = 24.$

Figure 4.10

Since $|\mathbf{B}| \cos\theta$ is the projection of $\mathbf{B}$ on $\mathbf{A}$, we can write

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}| \text{ times (projection of } \mathbf{B} \text{ on } \mathbf{A}), \tag{4.4}$$

or, alternatively,

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{B}| \text{ times (projection of } \mathbf{A} \text{ on } \mathbf{B}).$$

Also we find from (4.2) that

$$\mathbf{A} \cdot \mathbf{A} = |\mathbf{A}|^2 \cos 0° = |\mathbf{A}|^2 = A^2. \tag{4.5}$$

Sometimes $\mathbf{A}^2$ is written instead of $|\mathbf{A}|^2$ or A^2; you should understand that the square of a vector always means the square of its magnitude or its dot product with itself.

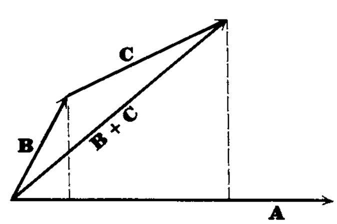


Figure 4.11

10344

From Figure 4.11 we can see that the projection of $\mathbf{B}+\mathbf{C}$ on $\mathbf{A}$ is equal to the projection of $\mathbf{B}$ on $\mathbf{A}$ plus the projection of $\mathbf{C}$ on $\mathbf{A}$. Then by (4.4)

$$\begin{aligned}(4.6)\quad \mathbf{A}\cdot(\mathbf{B}+\mathbf{C}) &= |\mathbf{A}| \text{ times (projection of } (\mathbf{B}+\mathbf{C}) \text{ on } \mathbf{A})\\ &= |\mathbf{A}| \text{ times (projection of } \mathbf{B} \text{ on } \mathbf{A} + \text{ projection of } \mathbf{C} \text{ on } \mathbf{A})\\ &= \mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{C}.\end{aligned}$$

This is the distributive law for scalar multiplication. By (4.3) we get also

$$(4.7)\qquad (\mathbf{B}+\mathbf{C})\cdot\mathbf{A} = \mathbf{B}\cdot\mathbf{A}+\mathbf{C}\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{C}.$$

The component form of $\mathbf{A}\cdot\mathbf{B}$ is very useful. We write

$$(4.8)\qquad \mathbf{A}\cdot\mathbf{B} = (\mathbf{i}A_x+\mathbf{j}A_y+\mathbf{k}A_z)\cdot(\mathbf{i}B_x+\mathbf{j}B_y+\mathbf{k}B_z).$$

By the distributive law we can multiply this out getting nine terms such as $A_xB_x\mathbf{i}\cdot\mathbf{i}$, $A_xB_y\mathbf{i}\cdot\mathbf{j}$, and so on. Using the definition of the scalar product, we find

$$(4.9)\qquad \begin{aligned}&\mathbf{i}\cdot\mathbf{i} = |\mathbf{i}|\cdot|\mathbf{i}|\cos 0^\circ = 1\cdot1\cdot1 = 1, \text{ and similarly, } \mathbf{j}\cdot\mathbf{j}=1,\ \mathbf{k}\cdot\mathbf{k}=1;\\ &\mathbf{i}\cdot\mathbf{j} = |\mathbf{i}|\cdot|\mathbf{j}|\cos 90^\circ = 1\cdot1\cdot0 = 0, \text{ and similarly, } \mathbf{i}\cdot\mathbf{k}=0,\ \mathbf{j}\cdot\mathbf{k}=0.\end{aligned}$$

Using (4.9) in (4.8), we get

$$(4.10)\qquad \boxed{\mathbf{A}\cdot\mathbf{B} = A_xB_x+A_yB_y+A_zB_z.}$$

Equation (4.10) is an important formula which you should memorize. There are several immediate uses of this formula and of the dot product.

Angle Between Two Vectors Given the vectors, we can find the angle between them by using both (4.2) and (4.10) and solving for $\cos\theta$.

▶ **Example 3.** Find the angle between the vectors $\mathbf{A}=3\mathbf{i}+6\mathbf{j}+9\mathbf{k}$ and $\mathbf{B}=-2\mathbf{i}+3\mathbf{j}+\mathbf{k}$. By (4.2) and (4.10) we get

$$(4.11)\qquad \begin{aligned}&\mathbf{A}\cdot\mathbf{B} = |\mathbf{A}|\,|\mathbf{B}|\cos\theta = 3\cdot(-2)+6\cdot3+9\cdot1 = 21,\\ &|\mathbf{A}| = \sqrt{3^2+6^2+9^2} = 3\sqrt{14},\quad |\mathbf{B}| = \sqrt{2^2+3^2+1^2} = \sqrt{14},\\ &3\sqrt{14}\sqrt{14}\cos\theta = 21,\quad \cos\theta = \frac{1}{2},\quad \theta = 60^\circ.\end{aligned}$$

Perpendicular and Parallel Vectors If two vectors are perpendicular, then $\cos\theta = 0$; thus

$$(4.12)\qquad \boxed{A_xB_x+A_yB_y+A_zB_z = 0 \quad \text{if } \mathbf{A} \text{ and } \mathbf{B} \text{ are perpendicular vectors.}}$$

If two vectors are parallel, their components are proportional; thus (when no components are zero)

FÊUE WHEP

$$(4.13) \qquad \frac{A_x}{B_x} = \frac{A_y}{B_y} = \frac{A_z}{B_z} \quad \text{if} \quad \mathbf{A} \text{ and } \mathbf{B} \text{ are parallel vectors.}$$

(Of course, if $B_x = 0$, then $A_x = 0$, etc.)

Vector Product The vector or cross product of **A** and **B** is written $\mathbf{A} \times \mathbf{B}$. By definition, $\mathbf{A} \times \mathbf{B}$ is a vector whose magnitude and direction are given as follows:

The magnitude of $\mathbf{A} \times \mathbf{B}$ is

$$(4.14) \qquad |\mathbf{A} \times \mathbf{B}| = |\mathbf{A}|\,|\mathbf{B}| \sin\theta,$$

where θ is the positive angle ($\leq 180°$) between **A** and **B**. The direction of $\mathbf{A} \times \mathbf{B}$ is perpendicular to the plane of **A** and **B** and in the sense **C** of advance of a right-handed screw rotated from **A** to **B** as in Figure 4.12.

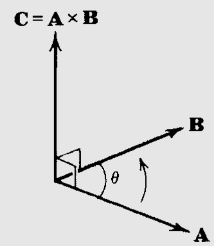


Figure 4.12

It is convenient to find the direction of $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ by the following right-hand rule. Think of grasping the line **C** (or a screwdriver driving a right-handed screw in the direction **C**) with the right hand. The fingers then curl in the direction of rotation of **A** into **B** (arrow in Figure 4.12) and the thumb points along $\mathbf{C} = \mathbf{A} \times \mathbf{B}$.

Perhaps the most startling result of the vector product definition is that $\mathbf{A} \times \mathbf{B}$ and $\mathbf{B} \times \mathbf{A}$ are not equal; in fact, $\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}$. In mathematical language, vector multiplication is not commutative.

We find from (4.14) that the cross product of any two parallel (or antiparallel) vectors has magnitude $|\mathbf{A} \times \mathbf{B}| = AB \sin 0° = 0$ (or $AB \sin 180° = 0$). Thus

$$(4.15) \qquad \begin{aligned} &\mathbf{A} \times \mathbf{B} = 0 \quad \text{if } \mathbf{A} \text{ and } \mathbf{B} \text{ are parallel or antiparallel,} \\ &\mathbf{A} \times \mathbf{A} = 0 \quad \text{for any } \mathbf{A}. \end{aligned}$$

Then we have the useful results

$$(4.16) \qquad \mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0.$$

Also from (4.14) we find

$$|\mathbf{i} \times \mathbf{j}| = |\mathbf{i}|\,|\mathbf{j}| \sin 90° = 1 \cdot 1 \cdot 1 = 1,$$

and similarly for the magnitude of the cross product of any two different unit vectors **i**, **j**, **k**. From the right-hand rule and Figure 4.13, we see that the direction of $\mathbf{i} \times \mathbf{j}$ is **k**, and since its magnitude is 1, we have $\mathbf{i} \times \mathbf{j} = \mathbf{k}$; however, $\mathbf{j} \times \mathbf{i} = -\mathbf{k}$. Similarly evaluating the other cross products, we find

10346

$$\text{(4.17)}\qquad \begin{array}{lll} \mathbf{i}\times\mathbf{j}=\mathbf{k} & \mathbf{j}\times\mathbf{k}=\mathbf{i} & \mathbf{k}\times\mathbf{i}=\mathbf{j}. \\ \mathbf{j}\times\mathbf{i}=-\mathbf{k} & \mathbf{k}\times\mathbf{j}=-\mathbf{i} & \mathbf{i}\times\mathbf{k}=-\mathbf{j}. \end{array}$$

A good way to remember these is to write them cyclically (around a circle as indicated in Figure 4.14). Reading around the circle counterclockwise (positive θ direction), we get the positive products (for example, $\mathbf{i}\times\mathbf{j}=\mathbf{k}$); reading the other way we get the negative products (for example, $\mathbf{i}\times\mathbf{k}=-\mathbf{j}$).

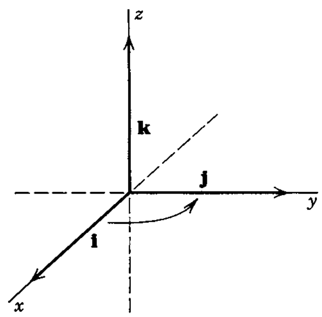


Figure 4.13

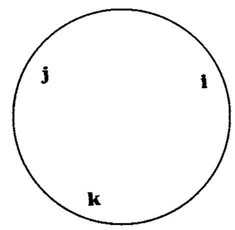


Figure 4.14

It is well to note here that the results (4.17) depend upon the way we have labeled the axes in Figure 4.13. We have arranged the (x, y, z) axes so that a rotation of the x into the y axis (through 90°) corresponds to the rotation of a right-handed screw advancing in the positive z direction. Such a coordinate system is called a *right-handed system*. If we used a left-handed system (say exchanging x and y), then all the equations in (4.17) would have their signs changed. This would be confusing; consequently, we practically always use right-handed coordinate systems, and we must be careful about this in drawing diagrams. (See Chapter 10, Section 6.)

To write $\mathbf{A}\times\mathbf{B}$ in component form we need the distributive law, namely

$$\text{(4.18)}\qquad \mathbf{A}\times(\mathbf{B}+\mathbf{C})=\mathbf{A}\times\mathbf{B}+\mathbf{A}\times\mathbf{C}.$$

(see Problem 7.18).

Then we find

$$\text{(4.19)}\qquad \begin{aligned} \mathbf{A}\times\mathbf{B} &= (\mathbf{i}A_x+\mathbf{j}A_y+\mathbf{k}A_z)\times(\mathbf{i}B_x+\mathbf{j}B_y+\mathbf{k}B_z) \\ &= \mathbf{i}(A_yB_z-A_zB_y)+\mathbf{j}(A_zB_x-A_xB_z)+\mathbf{k}(A_xB_y-A_yB_x) \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}. \end{aligned}$$

The second line in (4.19) is obtained by multiplying out the first line (getting nine products) and using (4.16) and (4.17). The determinant in (4.19) is the most convenient way to remember the component form of the vector product. You should verify that multiplying out the determinant using the elements of the first row gives the result in the line above it.

10347

Since $\mathbf{A} \times \mathbf{B}$ is a vector perpendicular to $\mathbf{A}$ and to $\mathbf{B}$, we can use (4.19) to find a vector perpendicular to two given vectors.

Example 4. Find a vector perpendicular to both $\mathbf{A} = 2\mathbf{i} + \mathbf{j} - \mathbf{k}$ and $\mathbf{B} = \mathbf{i} + 3\mathbf{j} - 2\mathbf{k}$.

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & -1 \\ 1 & 3 & -2 \end{vmatrix} = \mathbf{i}(-2+3) - \mathbf{j}(-4+1) + \mathbf{k}(6-1)$$
$$= \mathbf{i} + 3\mathbf{j} + 5\mathbf{k}.$$

PROBLEMS, SECTION 4

9. Let $\mathbf{A} = 2\mathbf{i} + 3\mathbf{j}$ and $\mathbf{B} = 4\mathbf{i} - 4\mathbf{j}$. Show graphically, and find algebraically, the vectors $-\mathbf{A}$, $3\mathbf{B}$, $\mathbf{A} - \mathbf{B}$, $\mathbf{B} + 2\mathbf{A}$, $\frac{1}{2}(\mathbf{A} + \mathbf{B})$.

10. If $\mathbf{A} + \mathbf{B} = 4\mathbf{j} - \mathbf{i}$ and $\mathbf{A} - \mathbf{B} = \mathbf{i} + 3\mathbf{j}$, find $\mathbf{A}$ and $\mathbf{B}$ algebraically. Show by a diagram how to find $\mathbf{A}$ and $\mathbf{B}$ geometrically.

11. Let $3\mathbf{i} - \mathbf{j} + 4\mathbf{k}$, $7\mathbf{j} - 2\mathbf{k}$, $\mathbf{i} - 3\mathbf{j} + \mathbf{k}$ be three vectors with tails at the origin. Then their heads determine three points A, B, C in space which form a triangle. Find vectors representing the sides AB, BC, CA in that order and direction (for example, A to B, not B to A) and show that the sum of these vectors is zero.

12. Find the angle between the vectors $\mathbf{A} = -2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$ and $\mathbf{B} = 2\mathbf{i} - 2\mathbf{j}$.

13. If $\mathbf{A} = 4\mathbf{i} - 3\mathbf{k}$ and $\mathbf{B} = -2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$, find the scalar projection of $\mathbf{A}$ on $\mathbf{B}$, the scalar projection of $\mathbf{B}$ on $\mathbf{A}$, and the cosine of the angle between $\mathbf{A}$ and $\mathbf{B}$.

14. Find the angles between (a) the space diagonals of a cube; (b) a space diagonal and an edge; (c) a space diagonal and a diagonal of a face.

15. Let $\mathbf{A} = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$. (a) Find a *unit* vector in the same direction as $\mathbf{A}$. *Hint:* Divide $\mathbf{A}$ by $|\mathbf{A}|$. (b) Find a vector in the same direction as $\mathbf{A}$ but of magnitude 12. (c) Find a vector perpendicular to $\mathbf{A}$. *Hint:* There are *many* such vectors; you are to find one of them. (d) Find a unit vector perpendicular to $\mathbf{A}$. See hint in (a).

16. Find a unit vector in the same direction as the vector $\mathbf{A} = 4\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$, and another unit vector in the same direction as $\mathbf{B} = -4\mathbf{i} + 3\mathbf{k}$. Show that the vector sum of these unit vectors bisects the angle between $\mathbf{A}$ and $\mathbf{B}$. *Hint:* Sketch the rhombus having the two unit vectors as adjacent sides.

17. Find three vectors (none of them parallel to a coordinate axis) which have lengths and directions such that they could be made into a right triangle.

18. Show that $2\mathbf{i} - \mathbf{j} + 4\mathbf{k}$ and $5\mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$ are orthogonal (perpendicular). Find a third vector perpendicular to both.

19. Find a vector perpendicular to both $\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}$ and $5\mathbf{i} - \mathbf{j} - 4\mathbf{k}$.

20. Find a vector perpendicular to both $\mathbf{i} + \mathbf{j}$ and $\mathbf{i} - 2\mathbf{k}$.

21. Show that $\mathbf{B}|\mathbf{A}| + \mathbf{A}|\mathbf{B}|$ and $\mathbf{A}|\mathbf{B}| - \mathbf{B}|\mathbf{A}|$ are orthogonal.

22. Square $(\mathbf{A} + \mathbf{B})$; interpret your result geometrically. *Hint:* Your answer is a law which you learned in trigonometry.

23. If $\mathbf{A} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$ and $\mathbf{A} \cdot \mathbf{B} = 0$, does it follow that $\mathbf{B} = 0$? (Either prove that it does or give a specific example to show that it doesn't.) Answer the same question if $\mathbf{A} \times \mathbf{B} = \mathbf{0}$. And again answer the same question if $\mathbf{A} \cdot \mathbf{B} = 0$ *and* $\mathbf{A} \times \mathbf{B} = \mathbf{0}$.

FEUE WHEP

24. What is the value of $(\mathbf{A} \times \mathbf{B})^2 + (\mathbf{A} \cdot \mathbf{B})^2$? *Comment*: This is a special case of Lagrange's identity. (*See* Chapter 6, Problem 3.12b, page 284.)

Use vectors as in Problems 3 to 8, and also the dot and cross product, to prove the following theorems from geometry.

25. The sum of the squares of the diagonals of a parallelogram is equal to twice the sum of the squares of two adjacent sides of the parallelogram.

26. The median to the base of an isosceles triangle is perpendicular to the base.

27. In a kite (four-sided figure made up of two pairs of equal adjacent sides), the diagonals are perpendicular.

28. The diagonals of a rhombus (four-sided figure with all sides of equal length) are perpendicular and bisect each other.

5. LINES AND PLANES

A great deal of analytic geometry can be simplified by the use of vector notation. Such things as equations of lines and planes, and distances between points or between lines and planes often occur in physics and it is very useful to be able to find them quickly. We shall talk about three-dimensional space most of the time although the ideas apply also to two dimensions. In analytic geometry a point is a set of three coordinates (x, y, z); we shall think of this point as the *head* of a vector $\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$ *with tail at the origin*. Most of the time the *vector* will be in the background of our minds and we shall not draw it; we shall just plot the point (x, y, z) which is the head of the vector. In other words, the point (x, y, z) and the vector $\mathbf{r}$ will be synonymous. We shall also use vectors joining two points. In Figure 5.1 the vector $\mathbf{A}$ from $(1, 2, 3)$ to (x, y, z) is

$$\mathbf{A} = \mathbf{r} - \mathbf{C} = (x, y, z) - (1, 2, 3) = (x - 1, y - 2, z - 3) \quad \text{or}$$
$$\mathbf{A} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z - (\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}) = \mathbf{i}(x - 1) + \mathbf{j}(y - 2) + \mathbf{k}(z - 3).$$

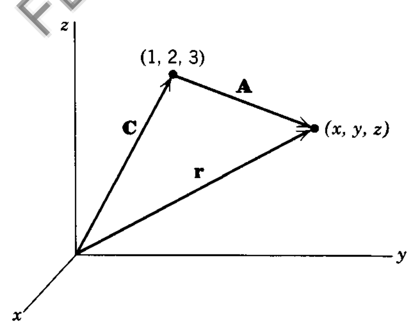


Figure 5.1

Thus we have two ways of writing vector equations; we may choose the one we prefer. Note the possible advantage of writing $(1, 0, -2)$ for $\mathbf{i} - 2\mathbf{k}$; since the zero is explicitly written, there is less chance of accidentally confusing $\mathbf{i} - 2\mathbf{k}$ with $\mathbf{i} - 2\mathbf{j} = (1, -2, 0)$. On the other hand, $5\mathbf{j}$ is simpler than $(0, 5, 0)$.

In two dimensions, we write the equation of a straight line through (x_0, y_0) with slope m as

$$\frac{y - y_0}{x - x_0} = m. \tag{5.1}$$

FEUE WHEP

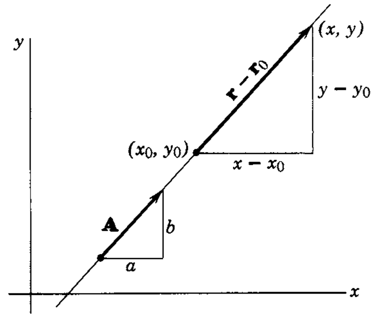


Figure 5.2

Suppose, instead of the slope, we are given a vector in the direction of the line, say $\mathbf{A} = \mathbf{i}a + \mathbf{j}b$ (Figure 5.2). Then the line through (x_0, y_0) and in the direction $\mathbf{A}$ is determined and we should be able to write its equation. The directed line segment from (x_0, y_0) to any point (x, y) on the line is the vector $\mathbf{r} - \mathbf{r}_0$ with components $x - x_0$ and $y - y_0$:

$$\mathbf{r} - \mathbf{r}_0 = \mathbf{i}(x - x_0) + \mathbf{j}(y - y_0). \tag{5.2}$$

This vector is parallel to $\mathbf{A} = \mathbf{i}a + \mathbf{j}b$. Now if two vectors are parallel, their components are proportional. Thus we can write (for $a, b \neq 0$)

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} \quad \text{or} \quad \frac{y - y_0}{x - x_0} = \frac{b}{a}. \tag{5.3}$$

This is the equation of the given straight line. As a check we see that the slope of the line is $m = b/a$, so (5.3) is the same as (5.1).

Another way to write this equation is to say that if $\mathbf{r} - \mathbf{r}_0$ and $\mathbf{A}$ are parallel vectors, one is some scalar multiple of the other, that is,

$$\mathbf{r} - \mathbf{r}_0 = \mathbf{A}t, \quad \text{or} \quad \mathbf{r} = \mathbf{r}_0 + \mathbf{A}t, \tag{5.4}$$

where t is the scalar multiple. We can think of t as a parameter; the component form of (5.4) is a set of parametric equations of the line, namely

$$\begin{aligned} x - x_0 &= at, \\ y - y_0 &= bt, \end{aligned} \quad \text{or} \quad \begin{aligned} x &= x_0 + at, \\ y &= y_0 + bt. \end{aligned} \tag{5.5}$$

Eliminating t yields the equation of the line in (5.3).

In three dimensions, the same ideas can be used. We want the equations of a straight line through a given point (x_0, y_0, z_0) and parallel to a given vector $\mathbf{A} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$. If (x, y, z) is any point on the line, the vector joining (x_0, y_0, z_0) and (x, y, z) is parallel to $\mathbf{A}$. Then its components $x - x_0$, $y - y_0$, $z - z_0$ are proportional to the components a, b, c of $\mathbf{A}$ and we have

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c} \qquad \text{(symmetric equations of a straight line, } a,\ b,\ c \neq 0). \tag{5.6}$$

If c, for instance, happens to be zero, we would have to write (5.6) in the form

FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP

(5.7)	$\dfrac{x-x_0}{a}=\dfrac{y-y_0}{b}, \quad z=z_0$	(symmetric equations of a straight line when $c=0$).

As in the two-dimensional case, equations (5.6) and (5.7) could be written

(5.8)	$\mathbf{r}=\mathbf{r}_0+\mathbf{A}t, \quad \text{or} \quad \begin{cases} x=x_0+at, \\ y=y_0+bt, \\ z=z_0+ct, \end{cases}$	(parametric equations of a straight line).

The parametric equations (5.8) have a particularly useful interpretation when the parameter t means time. Consider a particle m (electron, billiard ball, or star) moving along the straight line L in Figure 5.3. Position yourself at the origin and watch m move from P_0 to P along L. Your line of sight is the vector $\mathbf{r}$; it swings from $\mathbf{r}_0$ at $t=0$ to $\mathbf{r}=\mathbf{r}_0+\mathbf{A}t$ at time t. Note that the velocity of m is $d\mathbf{r}/dt=\mathbf{A}$; $\mathbf{A}$ is a vector along the line of motion.

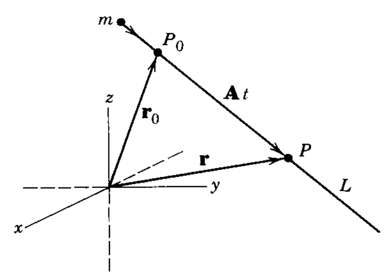


Figure 5.3

Going back to two dimensions, suppose we want the equation of a straight line L through the point (x_0,y_0) and perpendicular to a given vector $\mathbf{N}=a\mathbf{i}+b\mathbf{j}$. As above, the vector

$$\mathbf{r}-\mathbf{r}_0=(x-x_0)\mathbf{i}+(y-y_0)\mathbf{j}$$

lies along the line. This time we want this vector perpendicular to $\mathbf{N}$; recall that two vectors are perpendicular if their dot product is zero. Setting the dot product of $\mathbf{N}$ and $\mathbf{r}-\mathbf{r}_0$ equal to zero gives

$$(5.9) \qquad a(x-x_0)+b(y-y_0)=0 \quad \text{or} \quad \frac{y-y_0}{x-x_0}=-\frac{a}{b}.$$

This is the desired equation of the straight line L perpendicular to $\mathbf{N}$. As a check, note from Figure 5.4 that the slope of the line L is

$$\tan\theta=-\cot\phi=-a/b.$$

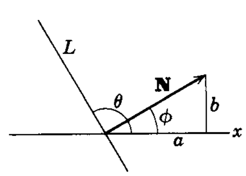


Figure 5.4

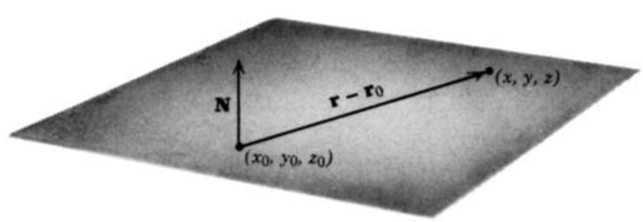


Figure 5.5

In three dimensions, we use this method to write the equation of a plane. If (x_0,y_0,z_0) is a given point in the plane and (x,y,z) is any other point in the plane,

FĚUE WHEP
10351

the vector (Figure 5.5)

$$\mathbf{r} - \mathbf{r}_0 = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}$$

is in the plane. If $\mathbf{N} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ is normal (perpendicular) to the plane, then $\mathbf{N}$ and $\mathbf{r} - \mathbf{r}_0$ are perpendicular, so the equation of the plane is $\mathbf{N} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$, or

$$(5.10) \qquad \begin{aligned} a(x - x_0) + b(y - y_0) + c(z - z_0) &= 0, \\ \text{or} \quad ax + by + cz &= d, \end{aligned} \qquad \text{(equation of a plane)}$$

where $d = ax_0 + by_0 + cz_0$.

If we are given equations like the ones above, we can read backwards to find $\mathbf{A}$ or $\mathbf{N}$. Thus we can say that the equations (5.6), (5.7), and (5.8) are the equations of a straight line which is parallel to the vector $\mathbf{A} = a\mathbf{i}+b\mathbf{j}+c\mathbf{k}$, and either equation in (5.10) is the equation of a plane perpendicular to the vector $\mathbf{N} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$.

Example 1. Find the equation of the plane through the three points $A(-1, 1, 1)$, $B(2, 3, 0)$, $C(0, 1, -2)$.

A vector joining any pair of the given points lies in the plane. Two such vectors are $\overrightarrow{AB} = (2, 3, 0) - (-1, 1, 1) = (3, 2, -1)$ and $\overrightarrow{AC} = (1, 0, -3)$. The cross product of these two vectors is perpendicular to the plane. This is

$$\mathbf{N} = (\overrightarrow{AB}) \times (\overrightarrow{AC}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 2 & -1 \\ 1 & 0 & -3 \end{vmatrix} = -6\mathbf{i} + 8\mathbf{j} - 2\mathbf{k}.$$

Now we write the equation of the plane with normal direction $\mathbf{N}$ through one of the given points, say B, using (5.10):

$$-6(x - 2) + 8(y - 3) - 2z = 0 \quad \text{or} \quad 3x - 4y + z + 6 = 0.$$

(Note that we could have divided $\mathbf{N}$ by -2 to save arithmetic.)

Example 2. Find the equations of a line through $(1, 0, -2)$ and perpendicular to the plane of Example 1.

The vector $3\mathbf{i}-4\mathbf{j}+\mathbf{k}$ is perpendicular to the plane of Example 1 and so parallel to the desired line. Thus by (5.6) the symmetric equations of the line are

$$\frac{(x - 1)}{3} = \frac{y}{-4} = \frac{(z + 2)}{1}.$$

By (5.8) the parametric equations of the line are $\mathbf{r} = \mathbf{i} - 2\mathbf{k} + (3\mathbf{i} - 4\mathbf{j} + \mathbf{k})t$ or, if you like, $\mathbf{r} = (1, 0, -2) + (3, -4, 1)t$.

Vectors give us a very convenient way of finding distances between points and lines or planes. Suppose we want to find the (perpendicular) distance from a point P

10352

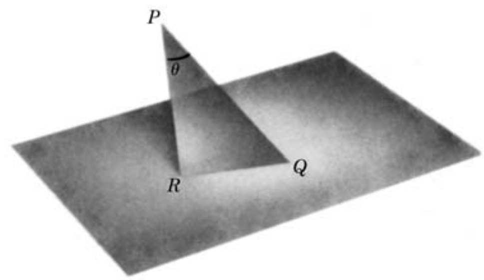


Figure 5.6

to the plane (5.10). (See Figure 5.6.) We pick *any* point Q we like in the plane (just by looking at the equation of the plane and thinking of some simple numbers x, y, z that satisfy it). The distance PR is what we want. Since PR and RQ are perpendicular (because PR is perpendicular to the plane), we have from Figure 5.6

$$PR = PQ\cos\theta. \tag{5.11}$$

From the equation of the plane, we can find a vector $\mathbf{N}$ normal to the plane. If we divide $\mathbf{N}$ by its magnitude, we have a unit vector normal to the plane; we denote this unit vector by $\mathbf{n}$. Then $|\overrightarrow{PQ}\cdot\mathbf{n}| = (PQ)\cos\theta$, which is what we need in (5.11) to find PR. (We have put in absolute value signs because $\overrightarrow{PQ}\cdot\mathbf{n}$ might be negative, whereas $(PQ)\cos\theta$, with θ acute as in Figure 5.6, is positive.)

▶ **Example 3.** Find the distance from the point $P(1,-2,3)$ to the plane $3x-2y+z+1=0$.

One point in the plane is $(1,2,0)$; call this point Q. Then the vector from P to Q is

$$\overrightarrow{PQ} = (1,2,0)-(1,-2,3) = (0,4,-3) = 4\mathbf{j}-3\mathbf{k}.$$

From the equation of the plane we get the normal vector

$$\mathbf{N} = 3\mathbf{i}-2\mathbf{j}+\mathbf{k}.$$

We get $\mathbf{n}$ by dividing $\mathbf{N}$ by $|\mathbf{N}| = \sqrt{14}$. Then we have

$$\begin{aligned}|PR| &= \left|\overrightarrow{PQ}\cdot\mathbf{n}\right| = \left|(4\mathbf{j}-3\mathbf{k})\cdot(3\mathbf{i}-2\mathbf{j}+\mathbf{k})/\sqrt{14}\right| \\ &= \left|(-8-3)/\sqrt{14}\right| = 11/\sqrt{14}.\end{aligned}$$

We can find the distance from a point P to a line in a similar way. In Figure 5.7 we want the perpendicular distance PR. We select any point on the line [that is, we pick any (x,y,z) satisfying the equations of the line]; call this point Q. Then (see Figure 5.7) $PR = PQ\sin\theta$. Let $\mathbf{A}$ be a vector along the line and $\mathbf{u}$ a unit vector along the line (obtained by dividing $\mathbf{A}$ by its magnitude). Then

$$\left|\overrightarrow{PQ}\times\mathbf{u}\right| = |PQ|\sin\theta,$$

so we get

$$|PR| = \left|\overrightarrow{PQ}\times\mathbf{u}\right|.$$

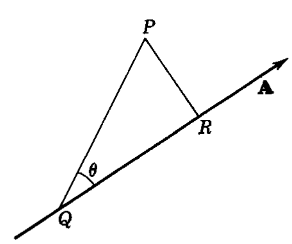


Figure 5.7

10353

Example 4. Find the distance from $P(1,2,-1)$ to the line joining $P_1(0,0,0)$ and $P_2(-1,0,2)$.

Let $\mathbf{A} = \overrightarrow{P_1P_2} = -\mathbf{i} + 2\mathbf{k}$; this is a vector along the line. Then a unit vector along the line is $\mathbf{u} = (1/\sqrt{5})(-\mathbf{i} + 2\mathbf{k})$. Let us take Q to be $P_1(0,0,0)$. Then $\overrightarrow{PQ} = -\mathbf{i} - 2\mathbf{j} + \mathbf{k}$, so we get for the distance $|PR|$:

$$|PR| = \frac{1}{\sqrt{5}}|(-\mathbf{i} - 2\mathbf{j} + \mathbf{k}) \times (-\mathbf{i} + 2\mathbf{k})| = \frac{1}{\sqrt{5}}|-4\mathbf{i} + \mathbf{j} - 2\mathbf{k}| = \sqrt{21/5}.$$

It is also straightforward to find the distance between two skew lines (and if you really want to appreciate vectors, just look up this calculation in an analytic geometry book that doesn't use vectors!). Pick two points P and Q, one on each line (Figure 5.8). Then $|\overrightarrow{PQ} \cdot \mathbf{n}|$, where $\mathbf{n}$ is a unit vector perpendicular to both lines, is the distance we want. Now if $\mathbf{A}$ and $\mathbf{B}$ are vectors along the two lines, then $\mathbf{A} \times \mathbf{B}$ is perpendicular to both, and $\mathbf{n}$ is just $\mathbf{A} \times \mathbf{B}$ divided by $|\mathbf{A} \times \mathbf{B}|$.

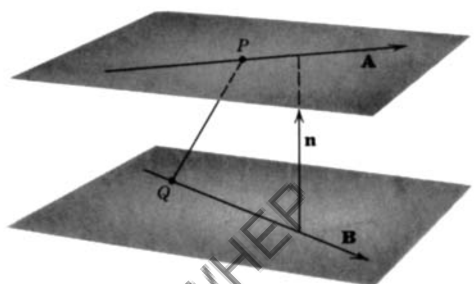


Figure 5.8

Example 5. Find the distance between the lines $\mathbf{r} = \mathbf{i} - 2\mathbf{j} + (\mathbf{i} - \mathbf{k})t$ and $\mathbf{r} = 2\mathbf{j} - \mathbf{k} + (\mathbf{j} - \mathbf{i})t$.

If we write the first line as $\mathbf{r} = \mathbf{r}_0 + \mathbf{A}t$, then (the head of) $\mathbf{r}_0$ is a simple choice for P, so we have

$$P = (1, -2, 0) \quad \text{and} \quad \mathbf{A} = \mathbf{i} - \mathbf{k}.$$

Similarly, from the second line we find

$$Q = (0, 2, -1) \quad \text{and} \quad \mathbf{B} = \mathbf{j} - \mathbf{i}.$$

Then $\mathbf{A} \times \mathbf{B} = \mathbf{i} + \mathbf{j} + \mathbf{k}$ and $\mathbf{n} = \left(1/\sqrt{3}\right)(\mathbf{i} + \mathbf{j} + \mathbf{k})$. Also

$$\overrightarrow{PQ} = (0, 2, -1) - (1, -2, 0) = (-1, 4, -1) = -\mathbf{i} + 4\mathbf{j} - \mathbf{k}.$$

Thus we get for the distance between the lines

$$\left|\overrightarrow{PQ} \cdot \mathbf{n}\right| = \left|(-\mathbf{i} + 4\mathbf{j} - \mathbf{k}) \cdot (\mathbf{i} + \mathbf{j} + \mathbf{k})/\sqrt{3}\right| = |-1 + 4 - 1|/\sqrt{3} = 2/\sqrt{3}.$$

Example 6. Find the direction of the line of intersection of the planes $x - 2y + 3z = 4$ and $2x + y - z = 5$.

The desired line lies in both planes, and so is perpendicular to the two normal vectors to the planes, namely $\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$ and $2\mathbf{i} + \mathbf{j} - \mathbf{k}$. Then the direction of the line is that of the cross product of these normal vectors; this is $-\mathbf{i} + 7\mathbf{j} + 5\mathbf{k}$.

10354

▶ **Example 7.** Find the cosine of the angle between the planes of Example 6.

The angle between the planes is the same as the angle between the normals to the planes. Thus our problem is to find the angle between the vectors $\mathbf{A} = \mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$ and $\mathbf{B} = 2\mathbf{i} + \mathbf{j} - \mathbf{k}$. Since $\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}|\,|\mathbf{B}| \cos\theta$, we have $-3 = \sqrt{14}\sqrt{6}\cos\theta$, and so $\cos\theta = -\sqrt{3/28}$. This gives the obtuse angle between the planes; the corresponding acute angle is $\pi - \theta$, or $\arccos\sqrt{3/28}$.

▶ PROBLEMS, SECTION 5

In Problems 1 to 5, all lines are in the (x, y) plane.

1. Write the equation of the straight line through $(2, -3)$ with slope $3/4$, in the parametric form $\mathbf{r} = \mathbf{r}_0 + \mathbf{A}t$.

2. Find the slope of the line whose parametric equation is $\mathbf{r} = (\mathbf{i} - \mathbf{j}) + (2\mathbf{i} + 3\mathbf{j})t$.

3. Write, in parametric form [as in Problem 1], the equation of the straight line that joins $(1, -2)$ and $(3, 0)$.

4. Write, in parametric form, the equation of the straight line that is perpendicular to $\mathbf{r} = (2\mathbf{i} + 4\mathbf{j}) + (\mathbf{i} - 2\mathbf{j})t$ and goes through $(1, 0)$.

5. Write, in parametric form, the equation of the y axis.

Find the symmetric equations (5.6) or (5.7) and the parametric equations (5.8) of a line, and/or the equation (5.10) of the plane satisfying the following given conditions.

6. Line through $(1, -1, -5)$ and $(2, -3, -3)$.

7. Line through $(2, 3, 4)$ and $(5, 1, -2)$.

8. Line through $(0, -2, 4)$ and $(3, -2, -1)$.

9. Line through $(-1, 3, 7)$ and $(-1, -2, 7)$.

10. Line through $(3, 4, -1)$ and parallel to $2\mathbf{i} - 3\mathbf{j} + 6\mathbf{k}$.

11. Line through $(4, -1, 3)$ and parallel to $\mathbf{i} - 2\mathbf{k}$.

12. Line through $(5, -4, 2)$ and parallel to the line $\mathbf{r} = \mathbf{i} - \mathbf{j} + (5\mathbf{i} - 2\mathbf{j} + \mathbf{k})t$.

13. Line through $(3, 0, -5)$ and parallel to the line $\mathbf{r} = (2, 1, -5) + (0, -3, 1)t$.

14. Plane containing the triangle ABC of Problem 4.11.

15. Plane through the origin and the points in Problem 8.

16. Plane through the point and perpendicular to the line in Problem 12.

17. Plane through the point and perpendicular to the line in Problem 13.

18. Plane containing the two parallel lines in Problem 12.

19. Plane containing the two parallel lines in Problem 13.

20. Plane containing the three points $(0, 1, 1)$, $(2, 1, 3)$, and $(4, 2, 1)$.

In Problems 21 to 23, find the angle between the given planes.

21. $2x + 6y - 3z = 10$ and $5x + 2y - z = 12$.

22. $2x - y - z = 4$ and $3x - 2y - 6z = 7$.

23. $2x + y - 2z = 3$ and $3x - 6y - 2z = 4$.

FEUE WHEP

24. Find a point on *both* the planes (that is, on their line of intersection) in Problem 21. Find a vector parallel to the line of intersection. Write the equations of the line of intersection of the planes. Find the distance from the origin to the line.

25. As in Problem 24, find the equations of the line of intersection of the planes in Problem 22. Find the distance from the point $(2, 1, -1)$ to the line.

26. As in Problem 24, find the equations of the line of intersection of the planes in Problem 23. Find the distance from the point $(1, 0, 0)$ to the line.

27. Find the equation of the plane through $(2, 3, -2)$ and perpendicular to both planes in Problem 21.

28. Find the equation of the plane through $(-4, -1, 2)$ and perpendicular to both planes in Problem 22.

29. Find a point on the plane $2x - y - z = 13$. Find the distance from $(7, 1, -2)$ to the plane.

30. Find the distance from the origin to the plane $3x - 2y - 6z = 7$.

31. Find the distance from $(-2, 4, 5)$ to the plane $2x + 6y - 3z = 10$.

32. Find the distance from $(3, -1, 2)$ to the plane $5x - y - z = 4$.

33. Find the perpendicular distance between the two parallel lines in Problem 12.

34. Find the distance (perpendicular is understood) between the two parallel lines in Problem 13.

35. Find the distance from $(2, 5, 1)$ to the line in Problem 10.

36. Find the distance from $(3, 2, 5)$ to the line in Problem 11.

37. Determine whether the lines

$$\frac{x-1}{2} = \frac{y+3}{1} = \frac{z-4}{-3} \quad \text{and} \quad \frac{x+3}{4} = \frac{y+4}{1} = \frac{8-z}{4}$$

intersect. *Two suggestions:* (1) Can you find the intersection point, if any? (2) Consider the distance between the lines.

38. Find the angle between the lines in Problem 37.

In Problems 39 and 40, show that the given lines intersect and find the acute angle between them.

39. $\mathbf{r} = 2\mathbf{j} + \mathbf{k} + (3\mathbf{i} - \mathbf{k})t_1 \quad \text{and} \quad \mathbf{r} = 7\mathbf{i} + 2\mathbf{k} + (2\mathbf{i} - \mathbf{j} + \mathbf{k})t_2.$

40. $\mathbf{r} = (5, -2, 0) + (1, -1, -1)t_1 \quad \text{and} \quad \mathbf{r} = (4, -4, -1) + (0, 3, 2)t_2.$

In Problems 41 to 44, find the distance between the two given lines.

41. $\mathbf{r} = (4, 3, -1) + (1, 1, 1)t \quad \text{and} \quad \mathbf{r} = (4, -1, 1) + (1, -2, -1)t.$

42. The line that joins $(0, 0, 0)$ to $(1, 2, -1)$, and the line that joins $(1, 1, 1)$ to $(2, 3, 4)$.

43. $\dfrac{x-1}{2} = \dfrac{y+2}{3} = \dfrac{2z-1}{4} \quad \text{and} \quad \dfrac{x+2}{-1} = \dfrac{2-y}{2}, \quad z = \dfrac{1}{2}.$

44. The x axis and $\mathbf{r} = \mathbf{j} - \mathbf{k} + (2\mathbf{i} - 3\mathbf{j} + \mathbf{k})t$.

45. A particle is traveling along the line $(x-3)/2 = (y+1)/(-2) = z - 1$. Write the equation of its path in the form $\mathbf{r} = \mathbf{r}_0 + \mathbf{A}t$. Find the distance of closest approach of the particle to the origin (that is, the distance from the origin to the line). If t represents time, show that the time of closest approach is $t = -(\mathbf{r}_0 \cdot \mathbf{A})/|\mathbf{A}|^2$. Use this value to check your answer for the distance of closest approach. *Hint:* See Figure 5.3. If P is the point of closest approach, what is $\mathbf{A} \cdot \mathbf{r}$?

FEUE WHEP
10356

▶ 6. MATRIX OPERATIONS

In Section 2 we used matrices simply as arrays of numbers. Now we want to go farther into the subject and discuss the meaning and use of multiplying a matrix by a number and of combining matrices by addition, subtraction, multiplication, and even (in a sense) division. We will see that we may be able to find functions of matrices such as e^{M}. These are, of course, all questions of definition, but we shall show some applications which might suggest reasonable definitions; or alternatively, given the definitions, we shall see what applications we can make of the matrix operations.

Matrix Equations Let us first emphasize again that two matrices are equal *only* if they are identical. Thus the matrix equation

$$\begin{pmatrix} x & r & u \\ y & s & v \end{pmatrix} = \begin{pmatrix} 2 & 1 & -5 \\ 3 & -7i & 1-i \end{pmatrix}$$

is really the set of six equations

$$x = 2, \qquad y = 3, \qquad r = 1, \qquad s = -7i, \qquad u = -5, \qquad v = 1 - i.$$

(Recall similar situations we have met before: The equation $z = x + iy = 2 - 3i$ is equivalent to the two real equations $x = 2$, $y = -3$; a vector equation in three dimensions is equivalent to three component equations.) In complicated problems involving many numbers or variables, it is often possible to save a great deal of writing by using a single matrix equation to replace a whole set of ordinary equations. Any time it is possible to so abbreviate the writing of a mathematical equation (like using a single letter for a complicated parenthesis) it not only saves time but often enables us to think more clearly.

Multiplication of a Matrix by a Number A convenient way to display the components of the vector $\mathbf{A} = 2\mathbf{i} + 3\mathbf{j}$ is to write them as elements of a matrix, either

$$\mathrm{A} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{called a column matrix or column vector,}$$

or

$$\mathrm{A}^{\mathrm{T}} = \begin{pmatrix} 2 & 3 \end{pmatrix} \quad \text{called a row matrix or row vector.}$$

The row matrix A^{T} is the transpose of the column matrix A. Observe the notation we are using here: We will often use the same letter for a vector and its column matrix, but we will usually write the letter representing the matrix as A (roman, not boldface), the vector as boldface $\mathbf{A}$, and the length of the vector as italic A.

Now suppose we want a vector of twice the length of $\mathbf{A}$ and in the same direction; we would write this as $2\mathbf{A} = 4\mathbf{i} + 6\mathbf{j}$. Then we would like to write its matrix representation as

$$2\mathrm{A} = 2\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}, \qquad 2\mathrm{A}^{\mathrm{T}} = 2(2 \quad 3) = (4 \quad 6).$$

This is, in fact, exactly how a matrix is multiplied by a number: *every* element of the matrix is multiplied by the number. Thus

$$k\begin{pmatrix} a & c & e \\ b & d & f \end{pmatrix} = \begin{pmatrix} ka & kc & ke \\ kb & kd & kf \end{pmatrix}$$

FÊUE WHEP
10357

and

$$\begin{pmatrix} -\frac{1}{2} & \frac{3}{4} \\ -1 & -\frac{5}{8} \end{pmatrix} = -\frac{1}{8}\begin{pmatrix} 4 & -6 \\ 8 & 5 \end{pmatrix}.$$

Note carefully a difference between determinants and matrices: multiplying a matrix by a number k means multiplying every element by k, but multiplying just *one* row of a determinant by k multiplies the determinant by k. Thus $\det(k\mathrm{A}) = k^2 \det \mathrm{A}$ for a 2 by 2 matrix, $\det(k\mathrm{A}) = k^3 \det \mathrm{A}$ for a 3 by 3 matrix, and so on.

Addition of Matrices When we add vectors algebraically, we add them by components. Matrices are added in the same way, by adding corresponding elements. For example,

$$\begin{aligned}(6.1)\qquad \begin{pmatrix} 1 & 3 & -2 \\ 4 & 7 & 1 \end{pmatrix} + \begin{pmatrix} 2 & -1 & 4 \\ 3 & -7 & -2 \end{pmatrix} &= \begin{pmatrix} 1+2 & 3-1 & -2+4 \\ 4+3 & 7-7 & 1-2 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 2 & 2 \\ 7 & 0 & -1 \end{pmatrix}.\end{aligned}$$

Note that if we add A + A we would get 2A in accord with our definition of twice a matrix above. Suppose we have

$$\mathrm{A} = \begin{pmatrix} 1 & 3 & -2 \\ 4 & 7 & 1 \end{pmatrix} \quad \text{and} \quad \mathrm{B} = \begin{pmatrix} 2 & -1 \\ 3 & 5 \end{pmatrix}.$$

In this case we cannot add A and B; we say that the sum is undefined or meaningless.

In applications, then, matrices are useful in representing things which are added by components. Suppose, for example, that, in (6.1), the columns represent displacements of three particles. The first particle is displaced by $\mathbf{i}+4\mathbf{j}$ (first column of the first matrix) and later by $2\mathbf{i}+3\mathbf{j}$ (first column of the second matrix). The total displacement is then $3\mathbf{i}+7\mathbf{j}$ (first column of the sum of the matrices). Similarly the second and third columns represent displacements of the second and third particles.

Multiplication of Matrices Let us start by defining the product of two matrices and then see what use we can make of the process. Here is a simple example to show what is meant by the product AB = C of two matrices A and B:

$$(6.2\text{a})\qquad \mathrm{AB} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix} = \mathrm{C}.$$

Observe that in the product matrix C, the element in the first row and first column is obtained by multiplying each element of the first row in A times the corresponding element in the first column of B and adding the results. This is referred to as "row times column" multiplication; when we compute $ae+bg$, we say that we have "multiplied the first row of A times the first column of B." Next examine the element $af+bh$ in the first row and second column of C; it is the "first row of A times the second column of B." Similarly, $ce+dg$ in the second row and first column of C is the "second row of A times the first column of B," and $cf+dh$ in the second

10358

row and second column of C is the "second row of A times the second column of B." Thus all the elements of C may be obtained by using the following simple rule:

(6.2b) The element in row i and column j of the product matrix AB is equal to row i of A times column j of B. In index notation

$$(\mathrm{AB})_{ij} = \sum_k A_{ik} B_{kj}.$$

Here is another useful way of saying this: Think of the elements in a row (or a column) of a matrix as the components of a vector. Then row times column multiplication for the matrix product AB corresponds to finding the dot product of a row vector of A and a column vector of B.

It is not necessary for matrices to be square in order for us to multiply them. Consider the following example.

▶ **Example 1.** Find the product of A and B if

$$\mathrm{A} = \begin{pmatrix} 4 & 2 \\ -3 & 1 \end{pmatrix}, \qquad \mathrm{B} = \begin{pmatrix} 1 & 5 & 3 \\ 2 & 7 & -4 \end{pmatrix}.$$

Following the rule we have stated, we get

$$\begin{aligned} \mathrm{AB} &= \begin{pmatrix} 4 & 2 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 & 3 \\ 2 & 7 & -4 \end{pmatrix} \\ &= \begin{pmatrix} 4 \cdot 1 + 2 \cdot 2 & 4 \cdot 5 + 2 \cdot 7 & 4 \cdot 3 + 2(-4) \\ -3 \cdot 1 + 1 \cdot 2 & -3 \cdot 5 + 1 \cdot 7 & -3 \cdot 3 + 1(-4) \end{pmatrix} \\ &= \begin{pmatrix} 8 & 34 & 4 \\ -1 & -8 & -13 \end{pmatrix}. \end{aligned}$$

Notice that the third column in B caused us no difficulty in following our rule; we simply multiplied each row of A times the third column of B to obtain the elements in the third column of AB. But suppose we tried to find the product BA. In B a row contains 3 elements, while in A a column contains only two; thus we are not able to apply the "row times column" method. Whenever this happens, we say that B is *not conformable* with respect to A, and the product BA is not defined (that is, it is meaningless and we do not use it).

The product AB (in that order) can be found if and only if the number of elements in a row of A equals the number of elements in a column of B; the matrices A, B in that order are then called *conformable*. (Observe that the number of rows in A and of columns in B have nothing to do with the question of whether we can find AB or not.)

FEUE WHEP

Example 2. Find AB and BA, given

$$A = \begin{pmatrix} 3 & -1 \\ -4 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 5 & 2 \\ -7 & 3 \end{pmatrix}.$$

Note that here the matrices are conformable in both orders, so we can find both AB and BA.

$$\begin{aligned} AB &= \begin{pmatrix} 3 & -1 \\ -4 & 2 \end{pmatrix}\begin{pmatrix} 5 & 2 \\ -7 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 3\cdot 5 - 1(-7) & 3\cdot 2 - 1\cdot 3 \\ -4\cdot 5 + 2(-7) & -4\cdot 2 + 2\cdot 3 \end{pmatrix} = \begin{pmatrix} 22 & 3 \\ -34 & -2 \end{pmatrix}. \\ BA &= \begin{pmatrix} 5 & 2 \\ -7 & 3 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -4 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 5\cdot 3 + 2(-4) & 5(-1) + 2\cdot 2 \\ -7\cdot 3 + 3(-4) & -7(-1) + 3\cdot 2 \end{pmatrix} = \begin{pmatrix} 7 & -1 \\ -33 & 13 \end{pmatrix}. \end{aligned}$$

Observe that AB is *not* the same as BA. We say that matrix multiplication is *not commutative*, or that, in general, matrices do not commute under multiplication. (Of course, two particular matrices may happen to commute.) We define the *commutator* of the matrices A and B by

(6.3) $$[A, B] = AB - BA = \textit{commutator} \text{ of A and B.}$$

(Commutators are of interest in classical and quantum mechanics.) Since matrices do not in general commute, be careful not to change the order of factors in a product of matrices unless you know they commute. For example

$$(A - B)(A + B) = A^2 + AB - BA - B^2 = A^2 - B^2 + [A, B].$$

This is not equal to $A^2 - B^2$ when A and B don't commute. Also see the discussion just after (6.17). On the other hand, the associative law is valid, that is, $A(BC) = (AB)C$, so we can write either as simply ABC. Also the distributive law holds: $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$ as we have been assuming above. (See Section 9.)

Zero Matrix The *zero* or *null* matrix means one with all its elements equal to zero. It is often abbreviated by 0, but we must be careful about this. For example:

(6.4) $$\text{If} \quad M = \begin{pmatrix} 2 & -4 \\ 1 & -2 \end{pmatrix}, \quad \text{then} \quad M^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

so we have $M^2 = 0$, but $M \neq 0$. Also see Problems 9 and 10.

FÊUE WHEP

Identity Matrix or Unit Matrix This is a square matrix with every element of the main diagonal (upper left to lower right) equal to 1 and all other elements equal to zero. For example

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.5}$$

is a unit or identity matrix of order 3 (that is, three rows and three columns). An identity or unit matrix is called 1 or I or U or E in various references. You should satisfy yourself that in multiplication, a unit matrix acts like the number 1, that is, if A is any matrix and I is the unit matrix conformable with A in the order in which we multiply, then $\mathrm{IA} = \mathrm{AI} = \mathrm{A}$ (Problem 11).

Operations with Determinants We do not define addition for determinants. However, multiplication is useful; we multiply determinants the same way we multiply matrices. It can be shown that if A and B are square matrices of the same order, then

$$\det \mathrm{AB} = \det \mathrm{BA} = (\det \mathrm{A}) \cdot (\det \mathrm{B}). \tag{6.6}$$

Look at Example 2 above to see that (6.6) is true even when matrices AB and BA are not equal, that is, when A and B do not commute.

Applications of Matrix Multiplication We can now write sets of simultaneous linear equations in a very simple form using matrices. Consider the matrix equation

$$\begin{pmatrix} 1 & 0 & -1 \\ -2 & 3 & 0 \\ 1 & -3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ -10 \end{pmatrix}. \tag{6.7}$$

If we multiply the first two matrices, we have

$$\begin{pmatrix} x - z \\ -2x + 3y \\ x - 3y + 2z \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ -10 \end{pmatrix}. \tag{6.8}$$

Now recall that two matrices are equal only if they are identical. Thus (6.8) is the set of three equations

$$\begin{cases} x \phantom{{}+3y} - z = 5 \\ -2x + 3y \phantom{{}+2z} = 1 \\ x - 3y + 2z = -10 \end{cases}. \tag{6.9}$$

Consequently (6.7) is the matrix form for the set of equations (6.9). In this way we can write any set of linear equations in matrix form. If we use letters to represent the matrices in (6.7),

$$\mathrm{M} = \begin{pmatrix} 1 & 0 & -1 \\ -2 & 3 & 0 \\ 1 & -3 & 2 \end{pmatrix}, \quad \mathrm{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \mathrm{k} = \begin{pmatrix} 5 \\ 1 \\ -10 \end{pmatrix}, \tag{6.10}$$

10361

then we can write (6.7) or (6.9) as

$$\text{(6.11)} \qquad \mathrm{Mr} = \mathrm{k}.$$

Or, in index notation, we can write $\sum_j M_{ij}x_j = k_i$. [Review Section 2, equations (2.3) to (2.6).] Note that (6.11) could represent any number of equations or unknowns (say 100 equations in 100 unknowns!). Thus we have a great simplification in notation which may help us to think more clearly about a problem. For example, if (6.11) were an ordinary algebraic equation, we would solve it for r to get

$$\text{(6.12)} \qquad \mathrm{r} = \mathrm{M}^{-1}\mathrm{k}.$$

Since M is a matrix, (6.12) only makes sense if we can give a meaning to M^{-1} such that (6.12) gives the solution of (6.7) or (6.9). Let's try to do this.

Inverse of a Matrix The reciprocal or inverse of a number x is x^{-1} such that the product $xx^{-1} = 1$. We define the inverse of a matrix M (if it has one) as the matrix M^{-1} such that MM^{-1} and $\mathrm{M}^{-1}\mathrm{M}$ are both equal to a unit matrix I. Note that only square matrices can have inverses (otherwise we could not multiply both MM^{-1} and $\mathrm{M}^{-1}\mathrm{M}$). Actually, some square matrices do not have inverses either. You can see from (6.6) that if $\mathrm{M}^{-1}\mathrm{M} = \mathrm{I}$, then $(\det \mathrm{M}^{-1})(\det \mathrm{M}) = \det \mathrm{I} = 1$. If two numbers have product $= 1$, then neither of them is zero; thus $\det \mathrm{M} \neq 0$ is a requirement for M to have an inverse.

If a matrix has an inverse we say that it is *invertible*; if it doesn't have an inverse, it is called *singular*. For simple numerical matrices your computer will easily produce the inverse of an invertible matrix. However, for theoretical purposes, we need a formula for the inverse; let's discuss this. The cofactor of an element in a square matrix M means exactly the same thing as the cofactor of that element in det M [see (3.3) and (3.4)]. Thus, the cofactor C_{ij} of the element m_{ij} in row i and column j is a number equal to $(-1)^{i+j}$ times the value of the determinant remaining when we cross off row i and column j. Then to find M^{-1}: Find the cofactors C_{ij} of all elements, write the matrix C whose elements are C_{ij}, transpose it (interchange rows and columns), and divide by det M. (See Problem 23.)

$$\text{(6.13)} \qquad \mathrm{M}^{-1} = \frac{1}{\det \mathrm{M}}\mathrm{C}^{\mathrm{T}} \quad \text{where } C_{ij} = \text{ cofactor of } m_{ij}$$

Although (6.13) is particularly useful in theoretical work, you should practice using it (as we said for Cramer's rule) on simple numerical problems in order to learn what the formula means.

FEUE WHEP
10362

▶ **Example 3.** For the matrix M of the coefficients in equations (6.7) or (6.9), find M^{-1}.

$$\mathrm{M} = \begin{pmatrix} 1 & 0 & -1 \\ -2 & 3 & 0 \\ 1 & -3 & 2 \end{pmatrix}.$$

We find $\det \mathrm{M} = 3$. The cofactors of the elements are:

$$1^{\text{st}} \text{ row}: \quad \begin{vmatrix} 3 & 0 \\ -3 & 2 \end{vmatrix} = 6, \quad -\begin{vmatrix} -2 & 0 \\ 1 & 2 \end{vmatrix} = 4, \quad \begin{vmatrix} -2 & 3 \\ 1 & -3 \end{vmatrix} = 3.$$

$$2^{\text{nd}} \text{ row}: \quad -\begin{vmatrix} 0 & -1 \\ -3 & 2 \end{vmatrix} = 3, \quad \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = 3, \quad -\begin{vmatrix} 1 & 0 \\ 1 & -3 \end{vmatrix} = 3.$$

$$3^{\text{rd}} \text{ row}: \quad \begin{vmatrix} 0 & -1 \\ 3 & 0 \end{vmatrix} = 3, \quad -\begin{vmatrix} 1 & -1 \\ -2 & 0 \end{vmatrix} = 2, \quad \begin{vmatrix} 1 & 0 \\ -2 & 3 \end{vmatrix} = 3.$$

Then

$$\mathrm{C} = \begin{pmatrix} 6 & 4 & 3 \\ 3 & 3 & 3 \\ 3 & 2 & 3 \end{pmatrix} \quad \text{so} \quad \mathrm{M}^{-1} = \frac{1}{\det \mathrm{M}} \mathrm{C}^{\mathrm{T}} = \frac{1}{3} \begin{pmatrix} 6 & 3 & 3 \\ 4 & 3 & 2 \\ 3 & 3 & 3 \end{pmatrix}.$$

Now we can use M^{-1} to solve equations (6.9). By (6.12), the solution is given by the column matrix $\mathrm{r} = \mathrm{M}^{-1}\mathrm{k}$, so we have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 6 & 3 & 3 \\ 4 & 3 & 2 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \\ -10 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -4 \end{pmatrix},$$

or $x = 1$, $y = 1$, $z = -4$. (See Problem 12.)

Rotation Matrices As another example of matrix multiplication, let's consider a case where we know the answer, just to see that our definition of matrix multiplication works the way we want it to. You probably know the rotation equations [for reference, see the next section, equation (7.12) and Figure 7.4]. Equation (7.12) gives the matrix which rotates the vector $\mathbf{r} = \mathbf{i}x + \mathbf{j}y$ through angle θ to become the vector $\mathbf{R} = \mathbf{i}X + \mathbf{j}Y$. Suppose we further rotate $\mathbf{R}$ through angle ϕ to become $\mathbf{R}' = \mathbf{i}X' + \mathbf{j}Y'$. We could write the matrix equations for the rotations in the form $\mathrm{R} = \mathrm{Mr}$ and $\mathrm{R}' = \mathrm{M}'\mathrm{R}$ where M and M′ are the rotation matrices (7.12) for rotation through angles θ and ϕ. Then, solving for R′ in terms of r, we get $\mathrm{R}' = \mathrm{M}'\mathrm{Mr}$. We expect the matrix product $\mathrm{M}'\mathrm{M}$ to give us the matrix for a rotation through the angle $\theta + \phi$, that is we expect to find

$$(6.14) \qquad \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}.$$

It is straightforward to multiply the two matrices (Problem 25) and verify (by using trigonometric identities) that (6.14) is correct. Also note that these two rotation matrices commute (that is, rotation through angle θ and then through angle ϕ gives the same result as rotation through ϕ followed by rotation through θ). This is true in this problem in two dimensions. As we will see in Section 7, rotation matrices in three dimensions do not in general commute if the two rotation axes are different. (See Problems 7.30 and 7.31.) But all rotations in the (x, y) plane are rotations about the z axis and so they commute.

FEUE WHEP

Functions of Matrices Since we now know how to multiply matrices and how to add them, we can evaluate any power of a matrix A and so evaluate a polynomial in A. The constant term c or $c\mathrm{A}^0$ in a polynomial is defined to mean c times the unit matrix I [see (6.16) below].

Example 4.

$$\text{If}\quad \mathrm{A}=\begin{pmatrix}1 & \sqrt{2}\\ -\sqrt{2} & -1\end{pmatrix},\quad \text{then}\quad \mathrm{A}^2=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}=-\mathrm{I},$$
$$\mathrm{A}^3=-\mathrm{A},\ \mathrm{A}^4=\mathrm{I},\quad \text{and so on.} \tag{6.15}$$

(Verify these powers and the fact that higher powers simply repeat these four results: A, −I, −A, I, over and over.) Then we can find (Problem 28)

$$\begin{aligned} f(\mathrm{A})&=3-2\mathrm{A}^2-\mathrm{A}^3-5\mathrm{A}^4+\mathrm{A}^6\\ &=3\mathrm{I}+2\mathrm{I}+\mathrm{A}-5\mathrm{I}-\mathrm{I}=\mathrm{A}-\mathrm{I}=\begin{pmatrix}0 & \sqrt{2}\\ -\sqrt{2} & -2\end{pmatrix}. \end{aligned} \tag{6.16}$$

We can extend this to other functions by expanding a given $f(x)$ in a power series if all the series we need to use happen to converge. For example, the series for e^z converges for all z, so we can find $e^{k\mathrm{A}}$ when A is a given matrix and k is any number, real or complex. Let A be the matrix in (6.15). Then (Problem 28), we find

$$\begin{aligned} e^{k\mathrm{A}}&=1+k\mathrm{A}+\frac{k^2\mathrm{A}^2}{2!}+\frac{k^3\mathrm{A}^3}{3!}+\frac{k^4\mathrm{A}^4}{4!}+\frac{k^5\mathrm{A}^5}{5!}+\cdots\\ &=(1-\frac{k^2}{2!}+\frac{k^4}{4!}+\cdots)\mathrm{I}+(k-\frac{k^3}{3!}+\frac{k^5}{5!})\mathrm{A}\\ &=(\cos k)\mathrm{I}+(\sin k)\mathrm{A}=\begin{pmatrix}\cos k+\sin k & \sqrt{2}\sin k\\ -\sqrt{2}\sin k & \cos k-\sin k\end{pmatrix}. \end{aligned} \tag{6.17}$$

A word of warning about functions of two matrices when A and B don't commute: Familiar formulas may mislead you; see (6.3) and the discussion following it. Be sure to write $(\mathrm{A}+\mathrm{B})^2=\mathrm{A}^2+\mathrm{AB}+\mathrm{BA}+\mathrm{B}^2$; don't write 2AB. Similarly, you can show that $e^{\mathrm{A}+\mathrm{B}}$ is not the same as $e^{\mathrm{A}}e^{\mathrm{B}}$ when A and B don't commute (see Problem 29 and Problem 15.34).

PROBLEMS, SECTION 6

In Problems 1 to 3, find AB, BA, A + B, A − B, A^2, B^2, 5A, 3B. Observe that $\mathrm{AB}\neq\mathrm{BA}$. Show that $(\mathrm{A}-\mathrm{B})(\mathrm{A}+\mathrm{B})\neq(\mathrm{A}+\mathrm{B})(\mathrm{A}-\mathrm{B})\neq\mathrm{A}^2-\mathrm{B}^2$. Show that $\det\mathrm{AB}=\det\mathrm{BA}=(\det\mathrm{A})(\det\mathrm{B})$, but that $\det(\mathrm{A}+\mathrm{B})\neq\det\mathrm{A}+\det\mathrm{B}$. Show that $\det(5\mathrm{A})\neq5\det\mathrm{A}$, and find n so that $\det(5\mathrm{A})=5^n\det\mathrm{A}$. Find similar results for det(3B). Remember that the point of doing these simple problems by hand is to learn how to manipulate determinants and matrices correctly. Check your answers by computer.

1. $\mathrm{A}=\begin{pmatrix}3 & 1\\ 2 & 5\end{pmatrix},\qquad \mathrm{B}=\begin{pmatrix}-2 & 2\\ 1 & 4\end{pmatrix}.$

2. $\mathrm{A}=\begin{pmatrix}2 & -5\\ -1 & 3\end{pmatrix},\qquad \mathrm{B}=\begin{pmatrix}-1 & 4\\ 0 & 2\end{pmatrix}.$

FEUE WHEP

3. $A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & -1 & 0 \\ 0 & 5 & 1 \end{pmatrix}$, $\quad B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 3 & -1 & 0 \end{pmatrix}$.

4. Given the matrices

$$A = \begin{pmatrix} 2 & 3 & 1 & -4 \\ 2 & 1 & 0 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 4 \\ 1 & -1 \\ 3 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 1 & 3 \\ 4 & -1 & -2 \\ -1 & 0 & 1 \end{pmatrix},$$

compute or mark as meaningless all products of two of these matrices (AB, BA, A^2, etc.); of three of them (ABC, A^2C, A^3, etc.).

5. Compute the product of each of the matrices in Problem 4 with its transpose [see (2.2) or (9.1)] in both orders, that is AA^T and A^TA, etc.

6. The Pauli spin matrices in quantum mechanics are

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(You will probably find these called σ_x, σ_y, σ_z in your quantum mechanics texts.) Show that $A^2 = B^2 = C^2 =$ a unit matrix. Also show that any two of these matrices anticommute, that is, $AB = -BA$, etc. Show that the commutator of A and B, that is, $AB - BA$, is $2iC$, and similarly for other pairs in cyclic order.

7. Find the matrix product

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

By evaluating this in two ways, verify the associative law for matrix multiplication, that is, $A(BC) = (AB)C$, which justifies our writing just ABC.

8. Show, by multiplying the matrices, that the following equation represents an ellipse.

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 5 & -7 \\ 7 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 30.$$

9. Find AB and BA given

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 10 & 4 \\ -5 & -2 \end{pmatrix}.$$

Observe that AB is the null matrix; if we call it 0, then $AB = 0$, but neither A nor B is 0. Show that A is singular.

10. Given

$$C = \begin{pmatrix} 7 & 6 \\ 2 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} -3 & 2 \\ 7 & 5 \end{pmatrix}$$

and A as in Problem 9, show that $AC = AD$, but $C \neq D$ and $A \neq 0$.

11. Show that the unit matrix I has the property that we associate with the number 1, that is, $IA = A$ and $AI = A$, assuming that the matrices are conformable.

12. For the matrices in Example 3, verify that MM^{-1} and $M^{-1}M$ both equal a unit matrix. Multiply $M^{-1}k$ to verify the solution of equations (6.9).

In Problems 13 to 16, use (6.13) to find the inverse of the given matrix.

13. $\begin{pmatrix} 6 & 9 \\ 3 & 5 \end{pmatrix}$ **14.** $\begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix}$

FEUE WHEP
10365

15. $\begin{pmatrix} -1 & 2 & 3 \\ 2 & 0 & -4 \\ -1 & -1 & 1 \end{pmatrix}$

16. $\begin{pmatrix} -2 & 0 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & 0 \end{pmatrix}$

17. Given the matrices

$$\mathrm{A} = \begin{pmatrix} 1 & -1 & 1 \\ 4 & 0 & -1 \\ 4 & -2 & 0 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}.$$

(a) Find A^{-1}, B^{-1}, $\mathrm{B}^{-1}\mathrm{AB}$, and $\mathrm{B}^{-1}\mathrm{A}^{-1}\mathrm{B}$.

(b) Show that the last two matrices are inverses, that is, that their product is the unit matrix.

18. Problem 17(b) is a special case of the general theorem that the inverse of a product of matrices is the product of the inverses in reverse order. Prove this. *Hint:* Multiply ABCD times $\mathrm{D}^{-1}\mathrm{C}^{-1}\mathrm{B}^{-1}\mathrm{A}^{-1}$ to show that you get a unit matrix.

In Problems 19 to 22, solve each set of equations by the method of finding the inverse of the coefficient matrix. *Hint:* See Example 3.

19. $\begin{cases} x - 2y = 5 \\ 3x + y = 15 \end{cases}$

20. $\begin{cases} 2x + 3y = -1 \\ 5x + 4y = 8 \end{cases}$

21. $\begin{cases} x + 2z = 8 \\ 2x - y = -5 \\ x + y + z = 4 \end{cases}$

22. $\begin{cases} x - y + z = 4 \\ 2x + y - z = -1 \\ 3x + 2y + 2z = 5 \end{cases}$

23. Verify formula (6.13). *Hint:* Consider the product of the matrices MC^{T}. Use Problem 3.8.

24. Use the method of solving simultaneous equations by finding the inverse of the matrix of coefficients, together with the formula (6.13) for the inverse of a matrix, to obtain Cramer's rule.

25. Verify (6.14) by multiplying the matrices and using trigonometric addition formulas.

26. In (6.14), let $\theta = \phi = \pi/2$ and verify the result numerically.

27. Do Problem 26 if $\theta = \pi/2$, $\phi = \pi/4$.

28. Verify the calculations in (6.15), (6.16), and (6.17).

29. Show that if A and B are matrices which don't commute, then $e^{\mathrm{A}+\mathrm{B}} \neq e^{\mathrm{A}}e^{\mathrm{B}}$, but if they do commute then the relation holds. *Hint:* Write out several terms of the infinite series for e^{A}, e^{B}, and $e^{\mathrm{A}+\mathrm{B}}$ and do the multiplications carefully assuming that A and B don't commute. Then see what happens if they do commute.

30. For the Pauli spin matrix A in Problem 6, find the matrices $\sin k\mathrm{A}$, $\cos k\mathrm{A}$, $e^{k\mathrm{A}}$, and $e^{ik\mathrm{A}}$ where $i = \sqrt{-1}$.

31. Repeat Problem 30 for the Pauli spin matrix C in Problem 6. *Hint*: Show that if a matrix is diagonal, say $\mathrm{D} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, then $f(\mathrm{D}) = \begin{pmatrix} f(a) & 0 \\ 0 & f(b) \end{pmatrix}$.

32. For the Pauli spin matrix B in Problem 6, find $e^{i\theta\mathrm{B}}$ and show that your result is a rotation matrix. Repeat the calculation for $e^{-i\theta\mathrm{B}}$.

FEUE WHEP 10366

▶ 7. LINEAR COMBINATIONS, LINEAR FUNCTIONS, LINEAR OPERATORS

Given two vectors $\mathbf{A}$ and $\mathbf{B}$, the vector $3\mathbf{A} - 2\mathbf{B}$ is called a "linear combination" of $\mathbf{A}$ and $\mathbf{B}$. In general, a linear combination of $\mathbf{A}$ and $\mathbf{B}$ means $a\mathbf{A} + b\mathbf{B}$ where a and b are scalars. Geometrically, if $\mathbf{A}$ and $\mathbf{B}$ have the same tail and do not lie along a line, then they determine a plane. You should satisfy yourself that all linear combinations of $\mathbf{A}$ and $\mathbf{B}$ then lie in the plane. It is also true that every vector in the plane can be written as a linear combination of $\mathbf{A}$ and $\mathbf{B}$; we shall consider this in Section 8. The vector $\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$ with tail at the origin (which we used in writing equations of lines and planes) is a linear combination of the unit basis vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$.

A function of a vector, say $f(\mathbf{r})$, is called linear if

(7.1) $$f(\mathbf{r}_1 + \mathbf{r}_2) = f(\mathbf{r}_1) + f(\mathbf{r}_2), \quad \text{and} \quad f(a\mathbf{r}) = af(\mathbf{r}),$$

where a is a scalar.

For example, if $\mathbf{A} = 2\mathbf{i} + 3\mathbf{j} - \mathbf{k}$ is a given vector, then $f(\mathbf{r}) = \mathbf{A} \cdot \mathbf{r} = 2x + 3y - z$ is a linear function because

$$f(\mathbf{r}_1 + \mathbf{r}_2) = \mathbf{A} \cdot (\mathbf{r}_1 + \mathbf{r}_2) = \mathbf{A} \cdot \mathbf{r}_1 + \mathbf{A} \cdot \mathbf{r}_2 = f(\mathbf{r}_1) + f(\mathbf{r}_2), \quad \text{and}$$
$$f(a\mathbf{r}) = \mathbf{A} \cdot (a\mathbf{r}) = a\mathbf{A} \cdot \mathbf{r} = af(\mathbf{r}).$$

On the other hand, $f(\mathbf{r}) = |\mathbf{r}|$ is *not* a linear function, because the length of the sum of two vectors is not in general the sum of their lengths. That is,

$$f(\mathbf{r}_1 + \mathbf{r}_2) = |\mathbf{r}_1 + \mathbf{r}_2| \neq |\mathbf{r}_1| + |\mathbf{r}_2| = f(\mathbf{r}_1) + f(\mathbf{r}_2),$$

as you can see from Figure 7.1. Also note that although we call $y = mx + b$ a linear equation (it is the equation of a straight line), the function $f(x) = mx + b$ is not linear (unless $b = 0$) because

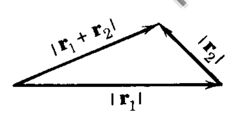


Figure 7.1

$$f(x_1 + x_2) = m(x_1 + x_2) + b \neq (mx_1 + b) + (mx_2 + b) = f(x_1) + f(x_2).$$

We can also consider vector functions of a vector $\mathbf{r}$. The magnetic field at each point (x, y, z), that is, at the head of the vector $\mathbf{r}$, is a vector $\mathbf{B} = \mathbf{i}B_x + \mathbf{j}B_y + \mathbf{k}B_z$. The components B_x, B_y, B_z may vary from point to pint, that is, they are functions of (x, y, z) or $\mathbf{r}$. Then

$\mathbf{F}(\mathbf{r})$ is a linear vector function if

(7.2) $$\mathbf{F}(\mathbf{r}_1 + \mathbf{r}_2) = \mathbf{F}(\mathbf{r}_1) + \mathbf{F}(\mathbf{r}_2) \quad \text{and} \quad \mathbf{F}(a\mathbf{r}) = a\mathbf{F}(\mathbf{r}),$$

where a is a scalar.

For example, $\mathbf{F}(\mathbf{r}) = b\mathbf{r}$ (where b is a scalar) is a linear vector function of $\mathbf{r}$.

FĒUE WHEP

You know from calculus that

$$\frac{d}{dx}[f(x)+g(x)] = \frac{d}{dx}f(x) + \frac{d}{dx}g(x) \quad \text{and} \qquad (7.3)$$
$$\frac{d}{dx}[kf(x)] = k\frac{d}{dx}f(x),$$

where k is a constant. We say that d/dx is a "linear operator" [compare (7.3) with (7.1) and (7.2)]. An "operator" or "operation" simply means a rule or some kind of instruction telling us what to do with whatever follows it. In other words, a linear operator is a linear function. Then

> O is a linear operator if
>
> $$O(A+B) = O(A) + O(B) \quad \text{and} \quad O(kA) = kO(A), \qquad (7.4)$$

where k is a number, and A and B are numbers, functions, vectors, and so on. Many of the errors people make happen because they assume that operators are linear when they are not (see problems).

Example 1. Is square root a linear operator? We are asking, is $\sqrt{A+B}$ the same as $\sqrt{A}+\sqrt{B}$? The answer is no; taking the square root is not a linear operation.

Example 2. Is taking the complex conjugate a linear operation? We want to know whether $\overline{A+B} = \bar{A}+\bar{B}$ and $\overline{kA} = k\bar{A}$. The first equation is true; the second equation is true if we restrict k to real numbers.

Matrix Operators, Linear Transformations Consider the set of equations

$$\begin{cases} X = ax + by, \\ Y = cx + dy, \end{cases} \quad \text{or} \quad \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{or} \quad \mathrm{R} = \mathrm{Mr}, \qquad (7.5)$$

where a, b, c, d, are constants. For every point (x, y), these equations give us a point (X, Y). If we think of each point of the (x, y) plane being moved to some other point (with some points like the origin not being moved), we can call this process a *mapping* or *transformation* of the plane into itself. All the information about this transformation is contained in the matrix M. We say that this matrix is an operator which maps the plane into itself. Any matrix can be thought of as an operator on (conformable) column matrices r. Since

$$\mathrm{M}(\mathrm{r}_1 + \mathrm{r}_2) = \mathrm{Mr}_1 + \mathrm{Mr}_2 \quad \text{and} \quad \mathrm{M}(k\mathrm{r}) = k(\mathrm{Mr}), \qquad (7.6)$$

the matrix M is a linear operator.

10368

Equations (7.5) can be interpreted geometrically in two ways. In Figure 7.2, we have one set of coordinate axes and the vector **r** has been changed to the vector **R** by the transformation (7.5). In Figure 7.3, we have *two* sets of coordinate axes,

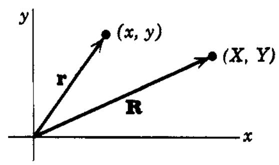


Figure 7.2

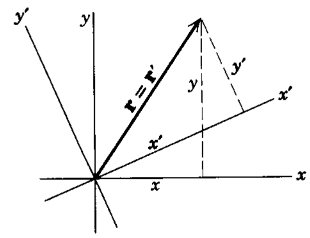


Figure 7.3

(x, y) and (x', y'), and *one* vector $\mathbf{r} = \mathbf{r}'$ with coordinates relative to each set of axes. This time the transformation

$$(7.7)\qquad \begin{cases} x' = ax + by, \\ y' = cx + dy, \end{cases} \quad \text{or} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{or} \quad \mathrm{r}' = \mathrm{M}\mathrm{r},$$

tells us how to get the components of the vector $\mathbf{r} = \mathbf{r}'$ relative to axes (x', y') when we know its components relative to axes (x, y).

Orthogonal Transformations We shall be particularly interested in the special case of a linear transformation which preserves the length of a vector. We call (7.7) an *orthogonal transformation* if

$$(7.8)\qquad x'^2 + y'^2 = x^2 + y^2,$$

and similarly for (7.5). You can see from the figures that this requirement says that the length of a vector is not changed by an orthogonal transformation. In Figure 7.2, the vector would be rotated (or perhaps reflected) with its length held fixed (that is $R = r$ for an orthogonal transformation). In Figure 7.3, the axes are rotated (or reflected), while the vector stays fixed. The matrix M of an orthogonal transformation is called an *orthogonal matrix.* Let's show that the inverse of an orthogonal matrix equals its transpose; in symbols

$$(7.9)\qquad \mathrm{M}^{-1} = \mathrm{M}^{\mathrm{T}}, \quad \mathrm{M} \text{ orthogonal.}$$

From (7.8) and (7.7) we have

$$\begin{aligned} x'^2 + y'^2 &= (ax + by)^2 + (cx + dy)^2 \\ &= (a^2 + c^2)x^2 + 2(ab + cd)xy + (b^2 + d^2)y^2 \equiv x^2 + y^2. \end{aligned}$$

Thus we must have $a^2 + c^2 = 1$, $b^2 + d^2 = 1$, $ab + cd = 0$. Then

$$(7.10)\qquad \begin{aligned} \mathrm{M}^{\mathrm{T}}\mathrm{M} &= \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

10369

Since M^TM is the unit matrix, M and M^T are inverse matrices as we claimed in (7.9). We have defined an orthogonal transformation in two dimensions and we have proved (7.9) for the 2-dimensional case. However, a square matrix of any order is called orthogonal if it satisfies (7.9), and you can easily show that the corresponding transformation preserves the lengths of vectors (Problem 9.24).

Now if we write (7.9) as $M^TM = I$ and use the facts from Section 3 that $\det(M^TM) = (\det M^T)(\det M)$ and $\det M^T = \det M$, we have $(\det M)^2 = \det(M^TM) = \det I = 1$, so

$$\det M = \pm 1, \quad \text{M orthogonal.} \tag{7.11}$$

This is true for M of any order since we have used only the definition (7.9) of an orthogonal matrix and some properties of determinants. As we shall see, $\det M = 1$ corresponds geometrically to a rotation, and $\det M = -1$ means that a reflection is involved.

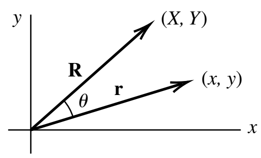


Figure 7.4

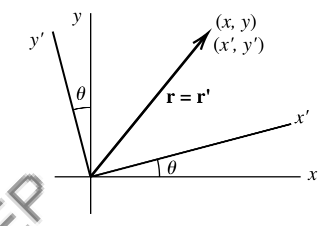


Figure 7.5

Rotations in 2 Dimensions In Figure 7.4, we have sketched the vector $\mathbf{r} = (x, y)$, and the vector $\mathbf{R} = (X, Y)$ which is the vector $\mathbf{r}$ rotated by angle θ. We write in matrix form the equations relating the components of $\mathbf{r}$ and $\mathbf{R}$ (Problem 19).

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{vector rotated.} \tag{7.12}$$

In Figure 7.5, we have sketched two sets of axes with the primed axes rotated by angle θ with respect to the unprimed axes. The vector $\mathbf{r} = (x, y)$, and the vector $\mathbf{r}' = (x', y')$ are the same vector, but with components relative to different axes. These components are related by the equations (Problem 20).

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{axes rotated.} \tag{7.13}$$

Both equations (7.12) and equations (7.13) are referred to as "rotation equations" and the θ matrices are called "rotation matrices". To distinguish them, we refer to the rotation (7.12) as an "active" transformation (vectors rotated), and to (7.13) as a "passive" transformation (vectors not moved but their components changed because the axes are rotated). Equations (7.7) or (7.13) are also referred to as a "change of basis". (Remember that we called **i**, **j**, **k** unit basis vectors; here we have changed from the **i**, **j**, **k** basis to the **i**′, **j**′, **k**′ basis. Also see Section 10.) Observe that the matrices in (7.12) and (7.13) are inverses of each other. You can see from the figures why this must be so. The rotation of a vector in, say, the counterclockwise direction produces the same result as the rotation of the axes in the opposite (clockwise) direction.

10370

We note that $\det M = \cos^2\theta + \sin^2\theta = 1$ for a rotation matrix. Any 2 by 2 orthogonal matrix with determinant 1 corresponds to a rotation, and any 2 by 2 orthogonal matrix with determinant $= -1$ corresponds to a reflection through a line.

▶ **Example 3.** Find what transformation corresponds to each of the following matrices.

$$\text{(7.14)}\qquad \mathrm{A} = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathrm{C} = \mathrm{AB}, \quad \mathrm{D} = \mathrm{BA}.$$

First we can show that all these matrices are orthogonal, and that $\det \mathrm{A} = 1$, but the determinants of the other three are -1 (Problem 21). Thus A is a rotation and B, C and D are reflections. Let's view these as active transformations (fixed axes, vectors rotated or reflected). Then by comparing A with (7.12), we have $\cos\theta = -1/2$, $\sin\theta = -\frac{1}{2}\sqrt{3}$, so this is a rotation of 240° (or −120°). Alternatively, we could ask what happens to the vector $\mathbf{i}$. We multiply matrix A times the column matrix $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and get

$$\frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ -\sqrt{3} \end{pmatrix} \quad \text{or} \quad -\frac{1}{2}(\mathbf{i} + \mathbf{j}\sqrt{3}),$$

which is $\mathbf{i}$ rotated by 240° as we had before.

Now B operating on $\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)$ leaves x fixed and changes the sign of y (check this); that is, B corresponds to a reflection through the x axis.

We find $\mathrm{C} = \mathrm{AB}$ and $\mathrm{D} = \mathrm{BA}$ by multiplying the matrices (Problem 21).

$$\text{(7.15)}\qquad \mathrm{C} = \mathrm{AB} = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}, \quad \mathrm{D} = \mathrm{BA} = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}.$$

We know that these are reflections since they have determinant $= -1$. To find the line through which the plane is reflected, we realize that the vectors along that line are unchanged by the reflection, so we want to find x and y, that is vector $\mathbf{r}$, which is mapped to itself by the transformation. For matrix C we write $\mathrm{Cr} = \mathrm{r}$.

$$\text{(7.16)}\qquad \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

You can verify (Problem 21) that the two equations in (7.16) are really the same equation, namely $y = -x\sqrt{3}$. Vectors along this line, say $\mathbf{i} - \mathbf{j}\sqrt{3}$, are not changed by the reflection [see (7.17)] so this is the reflection line. As further verification we can show [see (7.17)] that a vector perpendicular to this line, say $\mathbf{i}\sqrt{3} + \mathbf{j}$, is changed into its negative, that is, it is reflected through the line.

$$\text{(7.17)}\qquad \begin{aligned} \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix} &= \begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix}, \\ \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}\begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix} &= \begin{pmatrix} -\sqrt{3} \\ -1 \end{pmatrix}. \end{aligned}$$

Comment: The solution of the equation $\mathrm{Cr} = \mathrm{r}$ is an example of an eigenvalue, eigenvector problem. We shall discuss such problems in detail in Section 11.

We can analyze the transformation D in the same way we did C to find (Problem 21) that the reflection line is $y = x\sqrt{3}$. Note that matrices A and B do not commute and the transformations C and D are different.

FEUE WHEP

Rotations and Reflections in 3 Dimensions Let's consider 3 by 3 orthogonal matrices as active transformations rotating or reflecting vectors $\mathbf{r} = (x, y, z)$. A simple form for a rotation matrix is

$$\text{A} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.18}$$

You should satisfy yourself that this transformation produces a rotation of vectors about the z axis through angle θ. We can then find the rotation angle from (7.12) as we did in 2 dimensions. Similarly the matrix

$$\text{B} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{7.19}$$

produces a rotation about the z axis of angle θ together with a reflection through the (x, y) plane, and again we can find the rotation angle as in 2 dimensions.

We will show in Section 11 that any 3 by 3 orthogonal matrix with determinant $= 1$ can be written in the form (7.18) by choosing the z axis as the rotation axis, and any 3 by 3 orthogonal matrix with determinant $= -1$ can be written in the form (7.19). For now, let's look at a few simple problems we can do just by considering how the matrix maps certain vectors.

Example 4. The matrix for a rotation about the y axis is

$$\text{F} = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}. \tag{7.20}$$

You should satisfy yourself that the entry $-\sin\theta$ is in the right place for an active transformation. Let $\theta = 90°$; then the matrix F in (7.20) maps the vector $\mathbf{i} = (1, 0, 0)$ to the vector $-\mathbf{k} = (0, 0, -1)$; this is correct for a 90° rotation around the y axis. Check that $(0, 0, 1)$ is mapped to $(1, 0, 0)$.

Example 5. Find the mappings produced by the matrices

$$\text{G} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{K} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}. \tag{7.21}$$

First we find that the determinants are 1 so these are rotations. For G, either by inspection or by solving $\text{Gr} = \text{r}$ as in (7.16), we find that the vector $(1, 0, 1)$ is unchanged and so $\mathbf{i} + \mathbf{k}$ is the rotation axis. Now G^2 is the identity matrix (corresponding to a 360° rotation); thus the rotation angle for G is 180°.

Similarly for K, we find that the vector $(1, -1, 1)$ is unchanged by the transformation so $\mathbf{i} - \mathbf{j} + \mathbf{k}$ is the rotation axis. Now verify that K maps $\mathbf{i}$ to $-\mathbf{j}$, and $-\mathbf{j}$ to $\mathbf{k}$, and $\mathbf{k}$ to $\mathbf{i}$ (or, alternatively that K^3 is the identity matrix) so the rotation angle for K^3 is $\pm 360°$. From the geometry we see that the rotation $\mathbf{i} \to -\mathbf{j} \to \mathbf{k} \to \mathbf{i}$ is a rotation of $-120°$ about $\mathbf{i} - \mathbf{j} + \mathbf{k}$. (Also see Section 11.)

FEUE WHEP

▶ **Example 6.** Find the mapping produced by the matrix

$$\mathrm{L} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since $\det \mathrm{L} = -1$, this is a reflection through some plane. The vector perpendicular to the reflection plane is reversed by the reflection, so we ask for a vector satisfying $\mathrm{Lr} = -\mathrm{r}$. Either by solving these equations or by inspection we find $\mathbf{r} = (1, 1, 0) = \mathbf{i} + \mathbf{j}$. The reflecting plane is the plane through the origin perpendicular to this vector, that is, the plane $x + y = 0$ (see Section 5).

▶ PROBLEMS, SECTION 7

Are the following linear functions? Prove your conclusions by showing that $f(\mathbf{r})$ satisfies both of the equations (7.1) or that it does not satisfy at least one of them.

1. $f(\mathbf{r}) = \mathbf{A} \cdot \mathbf{r} + 3,$ where $\mathbf{A}$ is a given vector.

2. $f(\mathbf{r}) = \mathbf{A} \cdot (\mathbf{r} - \mathbf{k}z).$

3. $\mathbf{r} \cdot \mathbf{r}.$

Are the following linear vector functions? Prove your conclusions using (7.2).

4. $\mathbf{F}(\mathbf{r}) = \mathbf{r} - \mathbf{i}x = \mathbf{j}y + \mathbf{k}z.$

5. $\mathbf{F}(\mathbf{r}) = \mathbf{A} \times \mathbf{r},$ where $\mathbf{A}$ is a given vector.

6. $\mathbf{F}(\mathbf{r}) = \mathbf{r} + \mathbf{A},$ where $\mathbf{A}$ is a given vector.

Are the following operators linear?

7. Definite integral with respect to x from 0 to 1; the objects being operated on are functions of x.

8. Find the logarithm; operate on positive real numbers.

9. Find the square; operate on numbers or on functions.

10. Find the reciprocal; operate on numbers or on functions.

11. Find the absolute value; operate on complex numbers.

12. Let D stand for $\dfrac{d}{dx}$, D^2 for $\dfrac{d^2}{dx^2}$, $D^3 = \dfrac{d^3}{dx^3}$, and so on. Are D, D^2, D^3 linear? Operate on functions of x which can be differentiated as many times as needed.

13. (a) As in Problem 12, is $D^2 + 2D + 1$ linear?

(b) Is $x^2D^2 - 2xD + 7$ a linear operator?

14. Find the maximum; operate on functions of x.

15. Find the transpose; operate on matrices.

16. Find the inverse; operate on square matrices.

17. Find the determinant; operate on square matrices.

FEUE WHEP

18. With the cross product of two vectors defined by (4.14), show that finding the cross product is a linear operation, that is, show that (4.18) is valid. *Warning hint:* Don't try to prove it by writing out components: Writing, for example, $\mathbf{i}A_x \times (\mathbf{j}B_y + \mathbf{k}B_z) = \mathbf{i}A_x \times \mathbf{j}B_y + \mathbf{i}A_x \times \mathbf{k}B_z$ would be assuming what you're trying to prove. *Further hints:* First show that (4.18) is valid if $\mathbf{B}$ and $\mathbf{C}$ are both perpendicular to $\mathbf{A}$ by sketching (in the plane perpendicular to $\mathbf{A}$) the vectors $\mathbf{B}$, $\mathbf{C}$, $\mathbf{B}+\mathbf{C}$, and their vector products with $\mathbf{A}$. Then do the general case by first showing that $\mathbf{A}\times\mathbf{B}$ and $\mathbf{A}\times\mathbf{B}_\perp$ (where $\mathbf{B}_\perp$ is the vector component of $\mathbf{B}$ perpendicular to $\mathbf{A}$) have the same magnitude and the same direction.

19. If we multiply a complex number $z = re^{i\phi}$ by $e^{i\theta}$, we get $e^{i\theta}z = re^{i(\phi+\theta)}$, that is, a complex number with the same r but with its angle increased by θ. We can say that the vector $\mathbf{r}$ from the origin to the point $z = x + iy$ has been rotated by angle θ as in Figure 7.4 to become the vector $\mathbf{R}$ from the origin to the point $Z = X + iY$. Then we can write $X + iY = e^{i\theta}z = e^{i\theta}(x + iy)$. Take real and imaginary parts of this equation to obtain equations (7.12).

20. Verify equations (7.13) using Figure 7.5. *Hints*: Write $\mathbf{r}' = \mathbf{r}$ as $\mathbf{i}'x' + \mathbf{j}'y' = \mathbf{i}x + \mathbf{j}y$ and take the dot product of this equation with $\mathbf{i}'$ and with $\mathbf{j}'$ to get x' and y'. Evaluate the dot products of the unit vectors in terms of θ using Figure 7.5. For example, $\mathbf{i}'\cdot\mathbf{j}$ is the cosine of the angle between the x' axis and the y axis.

21. Do the details of Example 3 as follows:

(a) Verify that the four matrices in (7.14) are all orthogonal and verify the stated values of their determinants.

(b) Verify the products C = AB and D = BA in (7.15).

(c) Solve (7.16) to find the reflection line.

(d) Analyze the transformation D as we did C.

Let each of the following matrices represent an active transformation of vectors in the (x, y) plane (axes fixed, vectors rotated or reflected). As in Example 3, show that each matrix is orthogonal, find its determinant, and find the rotation angle, or find the line of reflection.

22. $\dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$

23. $\dfrac{1}{2}\begin{pmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{pmatrix}$

24. $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

25. $\dfrac{1}{3}\begin{pmatrix} -1 & 2\sqrt{2} \\ 2\sqrt{2} & 1 \end{pmatrix}$

26. $\dfrac{1}{5}\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$

27. $\dfrac{1}{\sqrt{2}}\begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}$

28. Write the matrices which produce a rotation θ about the x axis, or that rotation combined with a reflection through the (y, z) plane. [Compare (7.18) and (7.19) for rotation about the z axis.]

29. Construct the matrix corresponding to a rotation of 90° about the y axis together with a reflection through the (x, z) plane.

30. For the matrices G and K in (7.21), find the matrices R = GK and S = KG. Note that $\mathrm{R} \neq S$. (In 3 dimensions, rotations about two different axes do not in general commute.) Find what geometric transformations are produced by R and S.

FEUE WHEP

31. To see a physical example of non-commuting rotations, do the following experiment. Put a book on your desk and imagine a set of rectangular axes with the x and y axes in the plane of the desk with the z axis vertical. Place the book in the first quadrant with the x and y axes along the edges of the book. Rotate the book 90° about the x axis and then 90° about the z axis; note its position. Now repeat the experiment, this time rotating 90° about the z axis first, and then 90° about the x axis; note the different result. Write the matrices representing the 90° rotations and multiply them in both orders. In each case, find the axis and angle of rotation.

For each of the following matrices, find its determinant to see whether it produces a rotation or a reflection. If a rotation, find the axis and angle of rotation. If a reflection, find the reflecting plane and the rotation (if any) about the normal to this plane.

32. $\begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$ **33.** $\begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

34. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$ **35.** $\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

▶ 8. LINEAR DEPENDENCE AND INDEPENDENCE

We say that the three vectors $\mathbf{A} = \mathbf{i} + \mathbf{j}$, $\mathbf{B} = \mathbf{i} + \mathbf{k}$, and $\mathbf{C} = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$ are *linearly dependent* because $\mathbf{A} + \mathbf{B} - \mathbf{C} = \mathbf{0}$. The two vectors $\mathbf{i}$ and $\mathbf{j}$ are *linearly independent* because there are no numbers a and b (not *both* zero) such that the linear combination $a\mathbf{i} + b\mathbf{j}$ is zero. In general, a set of vectors is linearly dependent if some linear combination of them is zero (with not *all* the coefficients equal to zero). In the simple examples above, it was easy to see by inspection whether the vectors were linearly independent or not. In more complicated cases, we need a method of determining linear dependence. Consider the set of vectors

$$(1, 4, -5), (5, 2, 1), (2, -1, 3), \text{ and } (3, -6, 11); \tag{8.1}$$

We want to know whether they are linearly dependent, and if so, we want to find a smaller linearly independent set. Let us row reduce the matrix whose rows are the given vectors (see Section 2):

$$\begin{pmatrix} 1 & 4 & -5 \\ 5 & 2 & 1 \\ 2 & -1 & 3 \\ 3 & -6 & 11 \end{pmatrix} \to \begin{pmatrix} 9 & 0 & 7 \\ 0 & -9 & 13 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{8.2}$$

In row reduction, we are forming linear combinations of the rows by elementary row operations [see (2.8)]. All these operations are reversible, so we could, if we liked, reverse our calculations and combine the two vectors $(9, 0, 7)$ and $(0, -9, 13)$ to obtain each of the four original vectors (Problem 1). Thus there are only two independent vectors in (8.1); we refer to these independent vectors as basis vectors since all the original vectors can be written in terms of them (see Section 10). Note that the rank (see Section 2) of the matrix in (8.2) is equal to the number of independent or basis vectors.

FEUE WHEP

Linear Independence of Functions By a definition similar to that for vectors, we say that the functions $f_1(x)$, $f_2(x)$, $\cdots$, $f_n(x)$ are linearly dependent if some linear combination of them is identically zero, that is, if there are constants k_1, k_2, $\cdots$, k_n, not all zero, such that

$$k_1 f_1(x) + k_2 f_2(x) + \cdots + k_n f_n(x) \equiv 0. \tag{8.3}$$

For example, $\sin^2 x$ and $(1 - \cos^2 x)$ are linearly dependent since

$$\sin^2 x - (1 - \cos^2 x) \equiv 0.$$

But $\sin x$ and $\cos x$ are linearly independent since there are no numbers k_1 and k_2, not both zero, such that

$$k_1 \sin x + k_2 \cos x \tag{8.4}$$

is zero for *all* x (Problem 8).

We shall be particularly interested in knowing that a given set of functions is linearly independent. For this purpose the following theorem is useful (Problems 8 to 16, and Chapter 8, Section 5).

If $f_1(x)$, $f_2(x)$, $\cdots$, $f_n(x)$ have derivatives of order $n-1$, and if the determinant

$$W = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ f_1''(x) & f_2''(x) & \cdots & f_n''(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix} \not\equiv 0, \tag{8.5}$$

then the functions are linearly independent. (See Problem 16.) The determinant W is called the *Wronskian* of the functions.

Example 1. Using (8.5), show that the functions 1, x, $\sin x$ are linearly independent.

We write and evaluate the Wronskian,

$$W = \begin{vmatrix} 1 & x & \sin x \\ 0 & 1 & \cos x \\ 0 & 0 & -\sin x \end{vmatrix} = -\sin x.$$

Since $-\sin x$ is not identically equal to zero, the functions are linearly independent.

Example 2. Now let's compute the Wronskian for a case when the functions are linearly dependent.

$$W = \begin{vmatrix} x & \sin x & 2x - 3\sin x \\ 1 & \cos x & 2 - 3\cos x \\ 0 & -\sin x & 3\sin x \end{vmatrix} = \begin{vmatrix} x & \sin x & 2x \\ 1 & \cos x & 2 \\ 0 & -\sin x & 0 \end{vmatrix} = (\sin x)(2x - 2x) \equiv 0,$$

as we expected. However, note that "functions dependent" implies $W \equiv 0$, but $W \equiv 0$ does not necessarily imply "functions dependent". (See Problem 16.)

10376

Homogeneous Equations In Section 2 we considered sets of linear equations. Here we want to consider the special case of such equations when the constants on the right hand sides are all zero; these are called homogeneous equations. We write the homogeneous equations corresponding to (2.12) and (2.13) together with the row reduced matrices:

$$(8.6) \qquad \begin{cases} x + y = 0 \\ x - y = 0 \end{cases} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$(8.7) \qquad \begin{cases} x + y = 0 \\ 2x + 2y = 0 \end{cases} \qquad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We can draw several conclusions from these examples. Note that in (8.6) the only solution is $x = y = 0$; the rank of the matrix is 2, the same as the number of unknowns. In (8.7), the rank of the matrix is 1; this is less than the number of unknowns. This reflects what we could see in (8.7), that we really have just one equation in two unknowns; all the points on a line satisfy $x + y = 0$. In (8.8) we summarize the facts for homogeneous equations:

(8.8) Homogeneous equations are never inconsistent; they always have the solution "all unknowns = 0" (often called the "trivial solution"). If the number of independent equations (that is, the rank of the matrix) is the same as the number of unknowns, this is the only solution. If the rank of the matrix is less than the number of unknowns, there are infinitely many solutions.

A very important special case is a set of n homogeneous equations in n unknowns. By (8.8), these equations have only the trivial solution unless the rank of the matrix is less than n. This means that at least one row of the row reduced n by n matrix of the coefficients is a zero row. But then the determinant D of the coefficients is zero. Thus we have an important result (see Problems 21 to 25; also see Section 11):

(8.9) A system of n homogeneous equations in n unknowns has solutions other than the trivial solution if and only if the determinant of the coefficients is zero.

Solutions in Vector Form Geometrically, solutions of sets of linear equations may be points or lines or planes.

▶ **Example 3.** In Section 2, Example 4, we solved equations (2.15):

$$(8.10) \qquad x = 3 + 2z, \qquad y = 4 - z.$$

This solution set consists of all points on the line which is the intersection of these two planes. An interesting way to write the solution is the vector form

$$(8.11) \qquad \mathbf{r} = (x, y, z) = (3 + 2z, 4 - z, z) = (3, 4, 0) + (2, -1, 1)z.$$

FĒUE WHEP

If we put $z = t$, this is the parametric form of the equations of a straight line, $\mathbf{r} = \mathbf{r}_0 + \mathbf{A}t$ [see (5.8)].

Now let's consider the homogeneous equations (zero right hand sides) corresponding to equations (2.15). The equations and the row reduced matrix are:

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & -5 \\ -5 & 4 & 14 \\ 3 & -1 & -7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{8.12}$$

so the solutions are

$$x = 2z, \quad y = -z, \quad \text{or} \quad \mathbf{r} = (2, -1, 1)z. \tag{8.13}$$

Comparing (8.11) and (8.13), we see that the solution of the homogeneous equations Mr = 0 is a straight line through the origin; the solution of the equations Mr = k is a parallel straight line through the point $(3, 4, 0)$. We could say that the solution of Mr = k is the solution of the corresponding homogeneous equations plus the particular solution $\mathbf{r} = (3, 4, 0)$.

Here is an example of an important use of (8.9).

Example 4. For what values of λ does the following set of equations have nontrivial solutions for x and y? For each value of λ find the corresponding relation between x and y. This is an example of an *eigenvalue* problem; we shall discuss such problems in detail in Sections 11 and 12. The values of λ are called eigenvalues and the corresponding vectors (x, y) are called eigenvectors.

$$\begin{cases} (1-\lambda)x + 2y = 0, \\ 2x + (4-\lambda)y = 0. \end{cases} \tag{8.14}$$

By (8.9), we set the determinant M of the coefficients equal to zero. Then we solve for λ, and for each value of λ we solve for x and y.

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{vmatrix} = \lambda^2 - 5\lambda + 4 - 4 = \lambda(\lambda - 5) = 0, \quad \lambda = 0, 5.$$

For $\lambda = 0$, we find $x + 2y = 0$. For $\lambda = 5$, we find $2x - y = 0$. In vector notation the eigenvectors are: For $\lambda = 0$, $\mathbf{r} = (2, -1)s$, and for $\lambda = 5$, $\mathbf{r} = (1, 2)t$, where s and t are parameters in these vector equations of straight lines through the origin.

PROBLEMS, SECTION 8

1. Write each of the vectors (8.1) as a linear combination of the vectors $(9, 0, 7)$ and $(0, -9, 13)$. *Hint:* To get the right x component in $(1, 4, -5)$, you have to use $(1/9)(9, 0, 7)$. How do you get the right y component? Is the z component now correct?

In Problems 2 to 4, find out whether the given vectors are dependent or independent; if they are dependent, find a linearly independent subset. Write each of the given vectors as a linear combination of the independent vectors.

2. $(1, -2, 3)$, $(1, 1, 1)$, $(-2, 1, -4)$, $(3, 0, 5)$

FEUE WHEP
10378

3. $(0, 1, 1),\ (-1, 5, 3),\ (1, 0, 2),\ (2, -15, 1)$

4. $(3, 5, -1),\ (1, 4, 2),\ (-1, 0, 5),\ (6, 14, 5)$

5. Show that any vector $\mathbf{V}$ in a plane can be written as a linear combination of two non-parallel vectors $\mathbf{A}$ and $\mathbf{B}$ in the plane; that is, find a and b so that $\mathbf{V} = a\mathbf{A} + b\mathbf{B}$. *Hint:* Find the cross products $\mathbf{A} \times \mathbf{V}$ and $\mathbf{B} \times \mathbf{V}$; what are $\mathbf{A} \times \mathbf{A}$ and $\mathbf{B} \times \mathbf{B}$? Take components perpendicular to the plane to show that

$$a = \frac{(\mathbf{B} \times \mathbf{V}) \cdot \mathbf{n}}{(\mathbf{B} \times \mathbf{A}) \cdot \mathbf{n}}$$

where $\mathbf{n}$ is normal to the plane, and a similar formula for b.

6. Use Problem 5 to write $\mathbf{V} = 3\mathbf{i} + 5\mathbf{j}$ as a linear combination of $\mathbf{A} = 2\mathbf{i} + \mathbf{j}$ and $\mathbf{B} = 3\mathbf{i} - 2\mathbf{j}$. Show that the formulas in Problem 5, written as a quotient of 2 by 2 determinants, are just the Cramer's rule solution of simultaneous equations for a and b.

7. As in Problem 6, write $\mathbf{V} = 4\mathbf{i} - 5\mathbf{j}$ in terms of the basis vectors $\mathbf{i} - 4\mathbf{j}$ and $5\mathbf{i} + 2\mathbf{j}$.

In Problems 8 to 15, use (8.5) to show that the given functions are linearly independent.

8. $\sin x,\ \cos x$

9. $e^{ix},\ \sin x$

10. $x,\ e^x,\ xe^x$

11. $\sin x,\ \cos x,\ x \sin x,\ x \cos x$

12. $1,\ x^2,\ x^4,\ x^6$

13. $\sin x,\ \sin 2x$

14. $e^{ix},\ e^{-ix}$

15. $e^x,\ e^{ix},\ \cosh x$

16. (a) Prove that if the Wronskian (8.5) is not identically zero, then the functions f_1, f_2, $\ldots$, f_n are linearly independent. Note that this is equivalent to proving that if the functions are linearly dependent, then W is identically zero. *Hints:* Suppose (8.3) were true; you want to find the k's. Differentiate (8.3) repeatedly until you have a set of n equations for the n unknown k's. Then use (8.9).

(b) In part (a) you proved that if $W \not\equiv 0$, then the functions are linearly independent. You might think that if $W \equiv 0$, the functions would be linearly dependent. This is not necessarily true; if $W \equiv 0$, the functions might be either dependent or independent. For example, consider the functions x^3 and $|x^3|$ on the interval $(-1, 1)$. Show that $W \equiv 0$, but the functions are not linearly dependent on $(-1, 1)$. (Sketch them.) On the other hand, they are linearly dependent (in fact identical) on $(0, 1)$.

In Problems 17 to 20, solve the sets of homogeneous equations by row reducing the matrix.

17. $$\begin{cases} x - 2y + 3z = 0 \\ x + 4y - 6z = 0 \\ 2x + 2y - 3z = 0 \end{cases}$$

18. $$\begin{cases} 2x + 3z = 0 \\ 4x + 2y + 5z = 0 \\ x - y + 2z = 0 \end{cases}$$

19. $$\begin{cases} 3x + y + 3z + 6w = 0 \\ 4x - 7y - 3z + 5w = 0 \\ x + 3y + 4z - 3w = 0 \\ 3x + 2z + 7w = 0 \end{cases}$$

20. $$\begin{cases} 2x - 3y + 5z = 0 \\ x + 2y - z = 0 \\ x - 5y + 6z = 0 \\ 4x + y + 3z = 0 \end{cases}$$

21. Find a condition for four points in space to lie in a plane. Your answer should be in the form a determinant which must be equal to zero. *Hint:* The equation of a plane is of the form $ax + by + cz = d$, where a, b, c, d are constants. The four points (x_1, y_1, z_1), (x_2, y_2, z_2), etc., are all to satisfy this equation. When can you find a, b, c, d not all zero?

10379

22. Find a condition for three lines in a plane to intersect in one point. *Hint:* See Problem 21. Write the equation of a line as $ax + by = c$. Assume that no two of the lines are parallel.

Using (8.9), find the values of λ such that the following equations have nontrivial solutions, and for each λ, solve the equations. (See Example 4.)

23. $$\begin{cases} (4-\lambda)x - 2y = 0 \\ -2x + (7-\lambda)y = 0 \end{cases}$$

24. $$\begin{cases} (6-\lambda)x + 3y = 0 \\ 3x - (2+\lambda)y = 0 \end{cases}$$

25. $$\begin{cases} -(1+\lambda)x + y + 3z = 0, \\ x + (2-\lambda)y = 0, \\ 3x + (2-\lambda)z = 0. \end{cases}$$

For each of the following, write the solution in vector form [see (8.11) and (8.13)].

26. $$\begin{cases} 2x - 3y + 5z = 3 \\ x + 2y - z = 5 \\ x - 5y + 6z = -2 \\ 4x + y + 3z = 13 \end{cases}$$

27. $$\begin{cases} x - y + 2z = 3 \\ -2x + 2y - z = 0 \\ 4x - 4y + 5z = 6 \end{cases}$$

28. $$\begin{cases} 2x + y - 5z = 7 \\ x - 2y = 1 \\ 3x - 5y - z = 4 \end{cases}$$

9. SPECIAL MATRICES AND FORMULAS

In this section we want to discuss various terms used in work with matrices, and prove some important formulas. First we list for reference needed definitions and facts about matrices.

There are several special matrices which are related to a given matrix A. We outline in (9.1) what these matrices are called, what notations are used for them, and how we get them from A.

(9.1)	*Name of Matrix*	*Notations for it*	*How to get it from* A
	Transpose of A, or A transpose	A^T or $\widetilde{A}$ or A' or A^t	Interchange rows and columns in A.
	Complex conjugate of A	$\bar{A}$ or A^*	Take the complex conjugate of each element.
	Transpose conjugate, Hermitian conjugate, adjoint (Problem 9), Hermitian adjoint.	$A^\dagger$ (A dagger)	Take the complex conjugate of each element and transpose.
	Inverse of A	A^{-1}	See Formula (6.13).

There is another set of names for special types of matrices. In (9.2), we list these and their definitions for reference.

10380

(9.2)	A matrix is called	if it satisfies the condition(s)
	real	$\mathrm{A} = \bar{\mathrm{A}}$
	symmetric	$\mathrm{A} = \mathrm{A}^{\mathrm{T}}$, A real (matrix = its transpose)
	skew-symmetric or antisymmetric	$\mathrm{A} = -\mathrm{A}^{\mathrm{T}}$, A real
	orthogonal	$\mathrm{A}^{-1} = \mathrm{A}^{\mathrm{T}}$, A real (inverse = transpose)
	pure imaginary	$\mathrm{A} = -\bar{\mathrm{A}}$
	Hermitian	$\mathrm{A} = \mathrm{A}^{\dagger}$ (matrix = its transpose conjugate)
	anti-Hermitian	$\mathrm{A} = -\mathrm{A}^{\dagger}$
	unitary	$\mathrm{A}^{-1} = \mathrm{A}^{\dagger}$ (inverse = transpose conjugate)
	normal	$\mathrm{A}\mathrm{A}^{\dagger} = \mathrm{A}^{\dagger}\mathrm{A}$ (A and $\mathrm{A}^{\dagger}$ commute)

Now let's consider some examples and proofs using these terms.

Index Notation We are going to need index notation in some of our work below, so for reference we restate the rule in (6.2b) for matrix multiplication.

$$(9.3) \qquad (\mathrm{AB})_{ij} = \sum_k A_{ik} B_{kj}.$$

Study carefully the index notation for "row times column" multiplication. To find the element in row i and column j of the product matrix AB, we multiply row i of A times column j of B. Note that the k's (the sum is over k) are next to each other in (9.3). If we should happen to have $\sum_k B_{kj} A_{ik}$, we should rewrite it as $\sum_k A_{ik} B_{kj}$ (with the k's next to each other) to recognize it as an element of the matrix AB (not BA). We will see an example of this in (9.10) below.

Kronecker δ The *Kronecker* δ is defined by

$$(9.4) \qquad \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

For example, $\delta_{11} = 1$, $\delta_{12} = 0$, $\delta_{22} = 1$, $\delta_{31} = 0$, and so on. In this notation a unit matrix is one whose elements are δ_{ij} and we can write

$$(9.5) \qquad \mathrm{I} = (\delta_{ij}).$$

FEUE WHEP

10381

(*Also see* Chapter 10, Section 5.) The Kronecker δ notation is useful for other purposes. For example, since (for positive integers m and n)

$$\int_{-\pi}^{\pi} \cos nx \cos mx \, dx = \begin{cases} \pi, & \text{if } m = n, \\ 0, & \text{if } m \neq n, \end{cases} \tag{9.6a}$$

we can write

$$\int_{-\pi}^{\pi} \cos nx \cos mx \, dx = \pi \cdot \delta_{nm}. \tag{9.6b}$$

This is the same as (9.6a) because $\delta_{nm} = 0$ if $m \neq n$, and $\delta_{nm} = 1$ if $m = n$.

Using the Kronecker δ, we can give a formal proof that for any matrix M and a conformable unit matrix I, the product of I and M is just M. Using index notation and equations (9.3) and (9.4), we have

$$(\mathrm{IM})_{ij} = \sum_k \delta_{ik} M_{kj} = M_{ij} \quad \text{or} \quad \mathrm{IM} = \mathrm{M} \tag{9.7}$$

since $\delta_{ik} = 0$ unless $k = i$.

More Useful Theorems Let's use index notation to prove the associative law for matrix multiplication, that is

$$\mathrm{A(BC)} = \mathrm{(AB)C} = \mathrm{ABC}. \tag{9.8}$$

First we write $(\mathrm{BC})_{kj} = \sum_l B_{kl} C_{lj}$. Then we have

$$\begin{aligned} [\mathrm{A(BC)}]_{ij} &= \sum_k A_{ik} (\mathrm{BC})_{kj} = \sum_k A_{ik} \sum_l B_{kl} C_{lj} \\ &= \sum_k \sum_l A_{ik} B_{kl} C_{lj} = (\mathrm{ABC})_{ij} \end{aligned} \tag{9.9}$$

which is the index notation for A(BC) = ABC as in (9.8). We can prove (AB)C = ABC in a similar way (Problem 1).

In formulas we may want the transpose of the product of two matrices. First note that $A^{\mathrm{T}}_{ik} = A_{ki}$ [see (2.1) or (9.1)]. Then

$$\begin{aligned} (\mathrm{AB})^{\mathrm{T}}_{ik} &= (\mathrm{AB})_{ki} = \sum_j A_{kj} B_{ji} = \sum_j A^{\mathrm{T}}_{jk} B^{\mathrm{T}}_{ij} \\ &= \sum_j B^{\mathrm{T}}_{ij} A^{\mathrm{T}}_{jk} = (\mathrm{B^T A^T})_{ik}, \quad \text{or,} \\ (\mathrm{AB})^{\mathrm{T}} &= \mathrm{B^T A^T}. \end{aligned} \tag{9.10}$$

The theorem applies to a product of any number of matrices (see Problem 8b). For example

$$(\mathrm{ABCD})^{\mathrm{T}} = \mathrm{D^T C^T B^T A^T}. \tag{9.11}$$

The transpose of a product of matrices is equal to the product of the transposes in reverse order.

FEUE WHEP
10382

A similar theorem is true for the inverse of a product (see Section 6, Problem 18).

(9.12) $$(\mathrm{ABCD})^{-1} = \mathrm{D}^{-1}\mathrm{C}^{-1}\mathrm{B}^{-1}\mathrm{A}^{-1}.$$

The inverse of a product of matrices is equal to the product of the inverses in reverse order.

Trace of a Matrix The *trace* (or *spur*) or a square matrix A (written Tr A) is the sum of the elements on the main diagonal. Thus the trace of a unit n by n matrix is n, and the trace of the matrix M in (6.10) is 6. It is a theorem that the trace of a product of matrices is not changed by permuting them in cyclic order. For example

(9.13) $$\mathrm{Tr(ABC)} = \mathrm{Tr(BCA)} = \mathrm{Tr(CAB)}.$$

We can prove this as follows:

$$\begin{aligned}\mathrm{Tr(ABC)} &= \sum_i (\mathrm{ABC})_{ii} = \sum_i \sum_j \sum_k A_{ij}B_{jk}C_{ki} \\ &= \sum_i \sum_j \sum_k B_{jk}C_{ki}A_{ij} = \mathrm{Tr(BCA)} \\ &= \sum_i \sum_j \sum_k C_{ki}A_{ij}B_{jk} = \mathrm{Tr(CAB)}.\end{aligned}$$

Warning: Tr(ABC) is *not* equal to Tr(ACB) in general.

Theorem: If H is a Hermitian matrix, then $\mathrm{U} = e^{i\mathrm{H}}$ is a unitary matrix. (This is an important relation in quantum mechanics.) By (9.2) we need to prove that $\mathrm{U}^\dagger = \mathrm{U}^{-1}$ if $\mathrm{H}^\dagger = \mathrm{H}$. First, $e^{i\mathrm{H}}e^{-i\mathrm{H}} = e^{i\mathrm{H}-i\mathrm{H}}$ since H commutes with itself—see Problem 6.29. But this is e^0 which is the unit matrix [see Section 6] so $\mathrm{U}^{-1} = e^{-i\mathrm{H}}$. To find $\mathrm{U}^\dagger = (e^{i\mathrm{H}})^\dagger$, we expand $\mathrm{U} = e^{i\mathrm{H}}$ in a power series to get $\mathrm{U} = \sum_k (i\mathrm{H})^k/k!$ and then take the transpose conjugate. To do this we just need to realize that the transpose of a sum of matrices is the sum of the transposes, and that the transpose of a power of a matrix, say $(\mathrm{M}^n)^\mathrm{T}$ is equal to $(\mathrm{M}^\mathrm{T})^n$ (Problem 9.21). Also recall from Chapter 2 that you find the complex conjugate of an expression by changing the signs of all the i's. This means that $(i\mathrm{H})^\dagger = -i\mathrm{H}^\dagger = -i\mathrm{H}$ since H is Hermitian. Then summing the series we get $\mathrm{U}^\dagger = e^{-i\mathrm{H}}$, which is just what we found for U^{-1} above. Thus $\mathrm{U}^\dagger = \mathrm{U}^{-1}$, so U is a unitary matrix. (Also see Problem 11.61.)

► PROBLEMS, SECTION 9

1. Use index notation as in (9.9) to prove the second part of the associative law for matrix multiplication: (AB)C = ABC.

2. Use index notation to prove the distributive law for matrix multiplication, namely: A(B + C) = AB + AC.

FEUE WHEP

3. Given the following matrix, find the transpose, the inverse, the complex conjugate, and the transpose conjugate of A. Verify that $AA^{-1} = A^{-1}A =$ the unit matrix.

$$A = \begin{pmatrix} 1 & 0 & 5i \\ -2i & 2 & 0 \\ 1 & 1+i & 0 \end{pmatrix},$$

4. Repeat Problem 3 given

$$A = \begin{pmatrix} 0 & 2i & -1 \\ -i & 2 & 0 \\ 3 & 0 & 0 \end{pmatrix}.$$

5. Show that the product AA^{T} is a symmetric matrix.

6. Give numerical examples of: a symmetric matrix; a skew-symmetric matrix; a real matrix; a pure imaginary matrix.

7. Write each of the items in the second column of (9.2) in index notation.

8. (a) Prove that $(AB)^\dagger = B^\dagger A^\dagger$. *Hint:* See (9.10).

(b) Verify (9.11), that is, show that (9.10) applies to a product of any number of matrices. *Hint:* Use (9.10) and (9.8).

9. In (9.1) we have defined the adjoint of a matrix as the transpose conjugate. This is the usual definition except in algebra where the adjoint is defined as the transposed matrix of cofactors [see (6.13)]. Show that the two definitions are the same for a unitary matrix with determinant $= +1$.

10. Show that if a matrix is orthogonal and its determinant is +1, then each element of the matrix is equal to its own cofactor. *Hint:* Use (6.13) and the definition of an orthogonal matrix.

11. Show that a real Hermitian matrix is symmetric. Show that a real unitary matrix is orthogonal. *Note:* Thus we see that Hermitian is the complex analogue of symmetric, and unitary is the complex analogue of orthogonal. (See Section 11.)

12. Show that the definition of a Hermitian matrix ($A = A^\dagger$) can be written $a_{ij} = \bar{a}_{ji}$ (that is, the diagonal elements are real and the other elements have the property that $a_{12} = \bar{a}_{21}$, etc.). Construct an example of a Hermitian matrix.

13. Show that the following matrix is a unitary matrix.

$$\begin{pmatrix} (1+i\sqrt{3})/4 & \dfrac{\sqrt{3}}{2\sqrt{2}}(1+i) \\ \dfrac{-\sqrt{3}}{2\sqrt{2}}(1+i) & (\sqrt{3}+i)/4 \end{pmatrix}$$

14. Use (9.11) and (9.12) to simplify $(AB^{\mathrm{T}}C)^{\mathrm{T}}$, $(C^{-1}MC)^{-1}$, $(AH)^{-1}(AHA^{-1})^3(HA^{-1})^{-1}$.

15. (a) Show that the Pauli spin matrices (Problem 6.6) are Hermitian.

(b) Show that the Pauli spin matrices satisfy the Jacobi identity $\big[A,[B,C]\big] + \big[B,[C,A]\big] + \big[C,[A,B]\big] = 0$ where $[A,B]$ is the commutator of A, B [*see* (6.3)].

(c) Generalize (b) to prove the Jacobi identity for any (conformable) matrices A, B, C. *Also see* Chapter 6, Problem 3.14.

16. Let $C_{ij} = (-1)^{i+j}M_{ij}$ be the cofactor of element a_{ij} in the determinant A. Show that the statement of Laplace's development and the statement of Problem 3.8 can be combined in the equations

$$\sum_j a_{ij}C_{kj} = \delta_{ik} \cdot \det A, \quad \text{or} \quad \sum_i a_{ij}C_{ik} = \delta_{jk} \cdot \det A.$$

10384

17. (a) Show that if A and B are symmetric, then AB is not symmetric unless A and B commute.

(b) Show that a product of orthogonal matrices is orthogonal.

(c) Show that if A and B are Hermitian, then AB is not Hermitian unless A and B commute.

(d) Show that a product of unitary matrices is unitary.

18. If A and B are symmetric matrices, show that their commutator is antisymmetric [see equation (6.3)].

19. (a) Prove that $\mathrm{Tr}(\mathrm{AB}) = \mathrm{Tr}(\mathrm{BA})$. *Hint:* See proof of (9.13).

(b) Construct matrices A, B, C for which $\mathrm{Tr}(\mathrm{ABC}) \neq \mathrm{Tr}(\mathrm{CBA})$, but verify that $\mathrm{Tr}(\mathrm{ABC}) = \mathrm{Tr}(\mathrm{CAB})$.

(c) If S is a symmetric matrix and A is an antisymmetric matrix, show that $\mathrm{Tr}(\mathrm{SA}) = 0$. *Hint:* Consider $\mathrm{Tr}(\mathrm{SA})^{\mathrm{T}}$ and prove that $\mathrm{Tr}(\mathrm{SA}) = -\mathrm{Tr}(\mathrm{SA})$.

20. Show that the determinant of a unitary matrix is a complex number with absolute value = 1. *Hint:* See proof of equation (7.11).

21. Show that the transpose of a sum of matrices is equal to the sum of the transposes. Also show that $(\mathrm{M}^n)^{\mathrm{T}} = (\mathrm{M}^{\mathrm{T}})^n$. *Hint:* Use (9.11) and (9.8).

22. Show that a unitary matrix is a normal matrix, that is, that it commutes with its transpose conjugate [see (9.2)]. Also show that orthogonal, symmetric, antisymmetric, Hermitian, and anti-Hermitian matrices are normal.

23. Show that the following matrices are Hermitian whether A is Hermitian or not: $\mathrm{AA}^\dagger$, $\mathrm{A} + \mathrm{A}^\dagger$, $i(\mathrm{A} - \mathrm{A}^\dagger)$.

24. Show that an orthogonal transformation preserves the length of vectors. *Hint:* If r is the column matrix of vector $\mathbf{r}$ [see (6.10)], write out $\mathrm{r}^{\mathrm{T}}\mathrm{r}$ to show that it is the square of the length of $\mathbf{r}$. Similarly $\mathrm{R}^{\mathrm{T}}\mathrm{R} = |\mathbf{R}|^2$ and you want to show that $|\mathbf{R}|^2 = |\mathbf{r}|^2$, that is, $\mathrm{R}^{\mathrm{T}}\mathrm{R} = \mathrm{r}^{\mathrm{T}}\mathrm{r}$ if $\mathrm{R} = \mathrm{Mr}$ and M is orthogonal. Use (9.11).

25. (a) Show that the inverse of an orthogonal matrix is orthogonal. *Hint:* Let $\mathrm{A} = \mathrm{O}^{-1}$; from (9.2), write the condition for O to be orthogonal and show that A satisfies it.

(b) Show that the inverse of a unitary matrix is unitary. See hint in (a).

(c) If H is Hermitian and U is unitary, show that $\mathrm{U}^{-1}\mathrm{HU}$ is Hermitian.

▶ 10. LINEAR VECTOR SPACES

We have used extensively the vector $\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$ to mean a vector from the origin to the point (x, y, z). There is a one-to-one correspondence between the vectors $\mathbf{r}$ and the points (x, y, z); the collection of all such points or all such vectors makes up the 3-dimensional space often called R_3 (R for real) or V_3 (V for vector) or E_3 (E for Euclidean). Similarly, we can consider a 2-dimensional space V_2 of vectors $\mathbf{r} = \mathbf{i}x + \mathbf{j}y$ or points (x, y) making up the (x, y) plane. V_2 might also mean *any* plane through the origin. And V_1 means all the vectors from the origin to points on some line through the origin.

We also use x, y, z to mean the variables or unknowns in a problem. Now applied problems often involve more than three variables. By extension of the idea of V_3, it is convenient to call an ordered set of n numbers a point or vector in the n-dimensional space V_n. For example, the 4-vectors of special relativity are ordered

FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP 10385

sets of four numbers; we say that space-time is 4-dimensional. A point of the *phase space* used in classical and quantum mechanics is an ordered set of six numbers, the three components of the position of a particle and the three components of its momentum; thus the phase space of a particle is the 6-dimensional space V_6.

In such cases, we can't represent the variables as coordinates of a point in *physical* space since physical space has only three dimensions. But it is convenient and customary to extend our geometrical *terminology* anyway. Thus we use the terms *variables* and *coordinates* interchangeably and speak, for example, of a "point in 5-dimensional space," meaning an ordered set of values of five variables, and similarly for any number of variables. In three dimensions, we think of the coordinates of a point as the components of a vector from the origin to the point. By analogy, we call an ordered set of five numbers a "vector in 5-dimensional space" or an ordered set of n numbers a "vector in n-dimensional space."

Much of the geometrical terminology which is familiar in two and three dimensions can be extended to problems in n dimensions (that is, n variables) by using the algebra which parallels the geometry. For example, the distance from the origin to the point (x, y, z) is $\sqrt{x^2+y^2+z^2}$. By analogy in a problem in the five variables x, y, z, u, v, we define the distance from the origin $(0,0,0,0,0)$ to the point (x, y, z, u, v) as $\sqrt{x^2+y^2+z^2+u^2+v^2}$. By using the algebra which goes with the geometry, we can easily extend such ideas as the length of a vector, the dot product of two vectors, and therefore the angle between the vectors and the idea of orthogonality, etc. We saw in Section 7, that an orthogonal transformation in two or three dimensions corresponds to a rotation. Thus we might say, in a problem in n variables, that a linear transformation (that is a linear change of variables) satisfying "sum of squares of new variables = sum of squares of old variables" [compare (7.8)] corresponds to a "*rotation in n-dimensional space*."

Example 1. Find the distance between the points $(3, 0, 5, -2, 1)$ and $(0, 1, -2, 3, 0)$.

Generalizing what we would do in three dimensions, we find $d^2 = (3-0)^2 + (0-1)^2 + (5+2)^2 + (-2-3)^2 + (1-0)^2 = 9+1+49+25+1 = 85$, $d = \sqrt{85}$.

If we start with several vectors, and find linear combinations of them in the algebraic way (by components), then we say that the original set of vectors and all their linear combinations form a *linear vector space* (or just *vector space* or *linear space* or *space*). Note that if $\mathbf{r}$ is one of our original vectors, then $\mathbf{r} - \mathbf{r} = \mathbf{0}$ is one of the linear combinations; thus the zero vector (that is, the origin) must be a point in every vector space. A line or plane not passing through the origin is not a vector space.

Subspace, Span, Basis, Dimension Suppose we start with the four vectors in (8.1). We showed in (8.2) that they are all linear combinations of the two vectors $(9, 0, 7)$ and $(0, -9, 13)$. Now two linearly independent vectors (remember their tails are at the origin) determine a plane; all linear combinations of the two vectors lie in the plane. [The plane we are talking about in this example is the plane through the three points $(9, 0, 7)$, $(0, -9, 13)$, and the origin.] Since all the vectors making up this plane V_2 are also part of 3-dimensional space V_3, we call V_2 a *subspace* of V_3. Similarly any line lying in this plane and passing through the origin is a subspace of V_2 and of V_3. We say that either the original four vectors or the two independent ones *span* the space V_2; a set of vectors spans a space if all the vectors

FEUE WHEP
10386

in the space can be written as linear combinations of the spanning set. A set of *linearly independent* vectors which span a vector space is called a *basis*. Here the vectors $(9, 0, 7)$ and $(0, -9, 13)$ are one possible choice as a basis for the space V_2; another choice would be any two of the original vectors since in (8.2) no two of the vectors are dependent.

The *dimension* of a vector space is equal to the number of basis vectors. Note that this statement implies (correctly—see Problem 8) that no matter how you pick the basis vectors for a given vector space, there will always be the same number of them. This number is the dimension of the space. In 3 dimensions, we have frequently used the unit basis vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ which can also be written as $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$. Then in, say 5 dimensions, a corresponding set of unit basis vectors would be $(1, 0, 0, 0, 0)$, $(0, 1, 0, 0, 0)$, $(0, 0, 1, 0, 0)$, $(0, 0, 0, 1, 0)$, $(0, 0, 0, 0, 1)$. You should satisfy yourself that these five vectors are linearly independent and span a 5 dimensional space.

▶ **Example 2.** Find the dimension of the space spanned by the following vectors, and a basis for the space: $(1, 0, 1, 5, -2)$, $(0, 1, 0, 6, -3)$, $(2, -1, 2, 4, 1)$, $(3, 0, 3, 15, -6)$.

We write the matrix whose rows are the components of the vectors and row reduce it to find that there are three linearly independent vectors: $(1, 0, 1, 5, 0)$, $(0, 1, 0, 6, 0)$, $(0, 0, 0, 0, 1)$. These three vectors are a basis for the space which is therefore 3-dimensional.

Inner Product, Norm, Orthogonality Recall from (4.10) that the scalar (or dot or inner) product of two vectors $\mathbf{A} = (A_1, A_2, A_3)$ and $\mathbf{B} = (B_1, B_2, B_3)$ is $A_1B_1 + A_2B_2 + A_3B_3 = \sum_{i=1}^{3} A_iB_i$. This is very easy to generalize to n dimensions. By definition, the inner product of two vectors in n dimensions is given by

$$\text{(10.1)} \qquad \mathbf{A}\cdot\mathbf{B} = (\text{Inner product of } \mathbf{A} \text{ and } \mathbf{B}) = \sum_{i=1}^{n} A_iB_i.$$

Similarly, generalizing (4.1), we can define the length or *norm* of a vector in n dimensions by the formula:

$$\text{(10.2)} \qquad A = \text{Norm of } \mathbf{A} = ||\mathbf{A}|| = \sqrt{\mathbf{A}\cdot\mathbf{A}} = \sqrt{\sum_{i=1}^{n} A_i^2}.$$

In 3 dimensions, we also write the scalar product as $AB\cos\theta$ [see (4.2)] so if two vectors are orthogonal (perpendicular) their scalar product is $AB\cos\pi/2 = 0$. We generalize this to n dimensions by saying that two vectors in n dimensions are orthogonal if their inner product is zero.

$$\text{(10.3)} \qquad \mathbf{A} \text{ and } \mathbf{B} \text{ are orthogonal if } \sum_{i=1}^{n} A_iB_i = 0.$$

FEUE WHEP

Schwarz Inequality In 2 or 3 dimensions we can find the angle between two vectors [see (4.11)] from the formula $\mathbf{A}\cdot\mathbf{B} = AB\cos\theta$. It is tempting to use the same formula in n dimensions, but before we do we should be sure that the resulting value of $\cos\theta$ will satisfy $|\cos\theta| \le 1$, that is

$$(10.4)\qquad |\mathbf{A}\cdot\mathbf{B}| \le AB, \quad \text{or} \quad \left|\sum_{i=1}^{n} A_iB_i\right| \le \sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}.$$

This is called the Schwarz inequality (for n-dimensional Euclidean space). We can prove it as follows. First note that if $\mathbf{B} = \mathbf{0}$, (10.4) just says $0 \le 0$ which is certainly true. For $\mathbf{B} \ne \mathbf{0}$, we consider the vector $\mathbf{C} = B\mathbf{A} - (\mathbf{A}\cdot\mathbf{B})\mathbf{B}/B$, and find $\mathbf{C}\cdot\mathbf{C}$. Now $\mathbf{C}\cdot\mathbf{C} = \sum C_i^2 \ge 0$, so we have

$$(10.5)\qquad \begin{aligned}\mathbf{C}\cdot\mathbf{C} &= B^2(\mathbf{A}\cdot\mathbf{A}) - 2B(\mathbf{A}\cdot\mathbf{B})(\mathbf{A}\cdot\mathbf{B})/B + (\mathbf{A}\cdot\mathbf{B})^2(\mathbf{B}\cdot\mathbf{B})/B^2\\ &= A^2B^2 - 2(\mathbf{A}\cdot\mathbf{B})^2 + (\mathbf{A}\cdot\mathbf{B})^2 = A^2B^2 - (\mathbf{A}\cdot\mathbf{B})^2 = C^2 \ge 0,\end{aligned}$$

which gives (10.4). Thus, if we like, we can define the cosine of the angle between two vectors in n dimensions by $\cos\theta = \mathbf{A}\cdot\mathbf{B}/(AB)$. Note that equality holds in Schwarz's inequality if and only if $\cos\theta = \pm 1$, that is, when $\mathbf{A}$ and $\mathbf{B}$ are parallel or antiparallel, say $\mathbf{B} = k\mathbf{A}$.

▶ **Example 3.** Find the cosine of the angle between each pair of the 3 basis vectors we found in Example 2.

By (10.2) we find that the norms of the first two basis vectors are $\sqrt{1+1+25} = \sqrt{27}$ and $\sqrt{1+36} = \sqrt{37}$. By (10.1), the inner product of these two vectors is $1\cdot 0 + 0\cdot 1 + 1\cdot 0 + 5\cdot 6 + 0\cdot 0 = 30$. Thus $\cos\theta = 30/(\sqrt{27\cdot 37}) \simeq 0.949$, which, we note, is < 1 as Schwarz's inequality says. The third basis vector in Example 2 is orthogonal to the other two since the inner products are zero, that is, $\cos\theta = 0$.

Orthonormal Basis; Gram-Schmidt Method We call a set of vectors *orthonormal* if they are all mutually *ortho*gonal (perpendicular), and each vector is *normal*ized (that is, its norm is one—it has unit length). For example, the vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, form an orthonormal set. If we have a set of basis vectors for a space, it is often convenient to take combinations of them to form an orthonormal basis. The Gram-Schmidt method is a systematic process for doing this. It is very simple in idea although the details of carrying it out can get messy. Suppose we have basis vectors $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$. Normalize $\mathbf{A}$ to get the first vector of a set of orthonormal basis vectors. To get a second basis vector, subtract from $\mathbf{B}$ its component along $\mathbf{A}$; what remains is orthogonal to $\mathbf{A}$. [See equation (4.4) and Figure 4.10.] Normalize this remainder to find the second vector of an orthonormal basis. Similarly, subtract from $\mathbf{C}$ its components along $\mathbf{A}$ and $\mathbf{B}$ to find a third vector orthogonal to both $\mathbf{A}$ and $\mathbf{B}$ and normalize this third vector. We now have 3 mutually orthogonal unit vectors; this is the desired set of orthonormal basis vectors. In a space of higher dimension, this process can be continued. (We will see a use for this method in Section 11; see degeneracy, pages 152–153.)

10388

▶ **Example 4.** Given the basis vectors $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, below, use the Gram-Schmidt method to find an orthonormal set of basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Following the outline above, we find

$$\begin{aligned}
&\mathbf{A} = (0,0,5,0); & \mathbf{e}_1 &= \mathbf{A}/A = (0,0,1,0);\\
&\mathbf{B} = (2,0,3,0); & \mathbf{B} - (\mathbf{e}_1\cdot\mathbf{B})\mathbf{e}_1 &= \mathbf{B} - 3\mathbf{e}_1 = (2,0,0,0);\\
& & \mathbf{e}_2 &= (1,0,0,0);\\
&\mathbf{C} = (7,1,-5,3); & \mathbf{C} - (\mathbf{e}_1\cdot\mathbf{C})\mathbf{e}_1 - (\mathbf{e}_2\cdot\mathbf{C})\mathbf{e}_2 &= \mathbf{C} - (-5)\mathbf{e}_1 - 7\mathbf{e}_2\\
& & &= (0,1,0,3);\\
& & \mathbf{e}_3 &= (0,1,0,3)/\sqrt{10}.
\end{aligned}$$

Complex Euclidean Space In applications it is useful to allow vector components to be complex. For example, in three dimensions we might consider vectors like $(5+2i, 3-i, 1+i)$. Let's go back and see what modifications are needed in this case. In (10.2), we want the quantity under the square root sign to be positive. To assure this, we replace the square of A_i by the absolute square of A_i, that is by $|A_i|^2 = A_i^* A_i$ where A_i^* is the complex conjugate of A_i (see Chapter 2). Similarly, in (10.1) and (10.3), we replace $A_i B_i$ by $A_i^* B_i$. Thus we define

$$\text{(Inner product of } \mathbf{A} \text{ and } \mathbf{B}) = \sum_{i=1}^{n} A_i^* B_i \tag{10.6}$$

$$\text{(Norm of } \mathbf{A}) = ||\mathbf{A}|| = \sqrt{\sum_{i=1}^{n} A_i^* A_i} \tag{10.7}$$

$$\mathbf{A} \text{ and } \mathbf{B} \text{ are orthogonal if } \sum_{i=1}^{n} A_i^* B_i = 0. \tag{10.8}$$

The Schwarz inequality becomes (see Problem 6)

$$\left|\sum_{i=1}^{n} A_i^* B_i\right| \le \sqrt{\sum_{i=1}^{n} A_i^* A_i}\sqrt{\sum_{i=1}^{n} B_i^* B_i}. \tag{10.9}$$

Note that we can write the inner product in matrix form. If A is a column matrix with elements A_i, then the transpose conjugate matrix $\mathrm{A}^\dagger$ is a row matrix with elements A_i^*. Using this notation we can write $\sum A_i^* B_i = \mathrm{A}^\dagger \mathrm{B}$ (Problem 9).

▶ **Example 5.** Given $\mathbf{A} = (3i, 1-i, 2+3i, 1+2i)$, $\mathbf{B} = (-1, 1+2i, 3-i, i)$, $\mathbf{C} = (4-2i, 2-i, 1, i-2)$, we find by (10.6) to (10.8):

$$\begin{aligned}
\text{(Inner product of } \mathbf{A} \text{ and } \mathbf{B}) &= (-3i)(-1) + (1+i)(1+2i)\\
&\quad + (2-3i)(3-i) + (1-2i)i = 4-4i.
\end{aligned}$$

FEUE WHEP

$$(\text{Norm of } \mathbf{A})^2 = (-3i)(3i) + (1+i)(1-i) + (2-3i)(2+3i) + (1-2i)(1+2i)$$
$$= 9 + 2 + 13 + 5 = 29, \quad ||\mathbf{A}|| = \sqrt{29}.$$
$$(\text{Norm of } \mathbf{B})^2 = 1 + 5 + 10 + 1 = 17, \quad ||\mathbf{B}|| = \sqrt{17}.$$

Note that $|4-4i| = 4\sqrt{2} < \sqrt{29}\sqrt{17}$ in accord with the Schwarz inequality (10.9).

$$(\text{Inner product of } \mathbf{B} \text{ and } \mathbf{C}) = (-1)(4-2i) + (1-2i)(2-i) + (3+i)(1)$$
$$+ (-i)(i-2) = -4 + 2i - 5i + 3 + i + 1 + 2i = 0.$$

Thus by (10.8), $\mathbf{B}$ and $\mathbf{C}$ are orthogonal.

PROBLEMS, SECTION 10

1. Find the distance between the points

(a) $(4, -1, 2, 7)$ and $(2, 3, 1, 9)$;

(b) $(-1, 5, -3, 2, 4)$ and $(2, 6, 2, 7, 6)$;

(c) $(5, -2, 3, 3, 1, 0)$ and $(0, 1, 5, 7, 2, 1)$.

2. For the given sets of vectors, find the dimension of the space spanned by them and a basis for this space.

(a) $(1, -1, 0, 0)$, $(0, -2, 5, 1)$, $(1, -3, 5, 1)$, $(2, -4, 5, 1)$;

(b) $(0, 1, 2, 0, 0, 4)$, $(1, 1, 3, 5, -3, 5)$, $(1, 0, 0, 5, 0, 1)$, $(-1, 1, 3, -5, -3, 3)$, $(0, 0, 1, 0, -3, 0)$;

(c) $(0, 10, -1, 1, 10)$, $(2, -2, -4, 0, -3)$, $(4, 2, 0, 4, 5)$, $(3, 2, 0, 3, 4)$, $(5, -4, 5, 6, 2)$.

3. (a) Find the cosines of the angles between pairs of vectors in Problem 2(a).

(b) Find two orthogonal vectors in Problem 2(b).

4. For each given set of basis vectors, use the Gram-Schmidt method to find an orthonormal set.

(a) $\mathbf{A} = (0, 2, 0, 0)$, $\mathbf{B} = (3, -4, 0, 0)$, $\mathbf{C} = (1, 2, 3, 4)$.

(b) $\mathbf{A} = (0, 0, 0, 7)$, $\mathbf{B} = (2, 0, 0, 5)$, $\mathbf{C} = (3, 1, 1, 4)$.

(c) $\mathbf{A} = (6, 0, 0, 0)$, $\mathbf{B} = (1, 0, 2, 0)$, $\mathbf{C} = (4, 1, 9, 2)$.

5. By (10.6) and (10.7), find the norms of $\mathbf{A}$ and $\mathbf{B}$ and the inner product of $\mathbf{A}$ and $\mathbf{B}$, and note that the Schwarz inequality (10.9) is satisfied:

(a) $\mathbf{A} = (3+i, 1, 2-i, -5i, i+1)$, $\mathbf{B} = (2i, 4-3i, 1+i, 3i, 1)$;

(b) $\mathbf{A} = (2, 2i-3, 1+i, 5i, i-2)$, $\mathbf{B} = (5i-2, 1, 3+i, 2i, 4)$.

6. Write out the proof of the Schwarz inequality (10.9) for a complex Euclidean space. *Hint:* Follow the proof of (10.4) in (10.5), replacing the definitions of norm and inner product in (10.1) and (10.2) by the definitions in (10.6) and (10.7). Remember that norms are real and ≥ 0.

7. Show that, in n-dimensional space, any $n+1$ vectors are linearly dependent. *Hint:* See Section 8.

8. Show that two different sets of basis vectors for the same vector space must contain the same number of vectors. *Hint:* Suppose a basis for a given vector space contains n vectors. Use Problem 7 to show that there cannot be more than n vectors in a basis for this space. Conversely, if there were a correct basis with less than n vectors, what can you say about the claimed n-vector basis?

10390

9. Write equations (10.6) to (10.9) in matrix form as discussed just after (10.9).

10. Prove that $||\mathbf{A} + \mathbf{B}|| \le ||\mathbf{A}|| + ||\mathbf{B}||$. This is called the triangle inequality; in two or three dimensions, it simply says that the length of one side of a triangle $\le$ sum of the lengths of the other 2 sides. *Hint:* To prove it in n-dimensional space, write the square of the desired inequality using (10.2) and also use the Schwarz inequality (10.4). Generalize the theorem to complex Euclidean space by using (10.7) and (10.9).

11. EIGENVALUES AND EIGENVECTORS; DIAGONALIZING MATRICES

We can give the following physical interpretation to Figure 7.2 and equations (7.5). Suppose the (x, y) plane is covered by an elastic membrane which can be stretched, shrunk, or rotated (with the origin fixed). Then any point (x, y) of the membrane becomes some point (X, Y) after the deformation, and we can say that the matrix M describes the deformation. Let us now ask whether there are any vectors such that $\mathbf{R} = \lambda\mathbf{r}$ where $\lambda =$ const. Such vectors are called *eigenvectors* (or *characteristic* vectors) of the transformation, and the values of λ are called the *eigenvalues* (or *characteristic* values) of the matrix M of the transformation.

Eigenvalues To illustrate finding eigenvalues, let's consider the transformation

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{11.1}$$

The eigenvector condition $\mathbf{R} = \lambda\mathbf{r}$ is, in matrix notation,

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix},$$

or written out in equation form:

$$\begin{aligned} 5x - 2y &= \lambda x, \\ -2x + 2y &= \lambda y, \end{aligned} \qquad \text{or} \qquad \begin{aligned} (5-\lambda)x - 2y &= 0, \\ -2x + (2-\lambda)y &= 0. \end{aligned} \tag{11.2}$$

These equations are homogeneous. Recall from (8.9) that a set of homogeneous equations has solutions other than $x = y = 0$ only if the determinant of the coefficients is zero. Thus we want

$$\begin{vmatrix} 5-\lambda & -2 \\ -2 & 2-\lambda \end{vmatrix} = 0. \tag{11.3}$$

This is called the *characteristic equation* of the matrix M, and the determinant in (11.3) is called the *secular determinant.*

> To obtain the characteristic equation of a matrix M, we subtract λ from the elements on the main diagonal of M, and then set the determinant of the resulting matrix equal to zero.

We solve (11.3) for λ to find the characteristic values of M:

$$\begin{aligned} &(5-\lambda)(2-\lambda) - 4 = \lambda^2 - 7\lambda + 6 = 0, \\ &\lambda = 1 \quad \text{or} \quad \lambda = 6. \end{aligned} \tag{11.4}$$

FEUE WHEP

Eigenvectors Substituting the λ values from (11.4) into (11.2), we get:

$$\begin{aligned} &2x - y = 0 \qquad \text{from either of the equations (11.2) when } \lambda = 1; \\ &x + 2y = 0 \qquad \text{from either of the equations (11.2) when } \lambda = 6. \end{aligned} \tag{11.5}$$

We were looking for vectors $\mathbf{r} = \mathbf{i}x + \mathbf{j}y$ such that the transformation (11.1) would give an $\mathbf{R}$ parallel to $\mathbf{r}$. What we have found is that *any* vector $\mathbf{r}$ with x and y components satisfying either of the equations (11.5) has this property. Since equations (11.5) are equations of straight lines through the origin, such vectors lie along these lines (Figure 11.1). Then equations (11.5) show that any vector $\mathbf{r}$ from the origin to a point on $x + 2y = 0$ is changed by the transformation (11.1) to a vector in the same direction but six times as long, and any vector from the origin to a point on $2x - y = 0$ is unchanged by the transformation (11.1). These vectors (along $x + 2y = 0$ and $2x - y = 0$) are the eigenvectors of the transformation. Along these two directions (and only these), the deformation of the elastic membrane was a pure stretch with no shear (rotation).

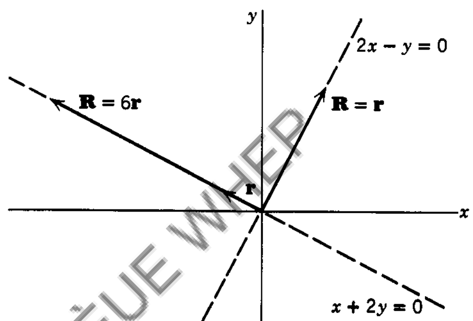


Figure 11.1

Diagonalizing a Matrix We next write (11.2) once with $\lambda = 1$, and again with $\lambda = 6$, using subscripts 1 and 2 to identify the corresponding eigenvectors:

$$\begin{aligned} 5x_1 - 2y_1 &= x_1, \qquad & 5x_2 - 2y_2 &= 6x_2, \\ -2x_1 + 2y_1 &= y_1, \qquad & -2x_2 + 2y_2 &= 6y_2. \end{aligned} \tag{11.6}$$

These four equations can be written as one matrix equation, as you can easily verify by multiplying out both sides (Problem 1):

$$\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}. \tag{11.7}$$

All we really can say about (x_1, y_1) is that $2x_1 - y_1 = 0$; however, it is convenient to pick numerical values of x_1 and y_1 to make $\mathbf{r}_1 = (x_1, y_1)$ a unit vector, and similarly for $\mathbf{r}_2 = (x_2, y_2)$. Then we have

$$x_1 = \frac{1}{\sqrt{5}}, \quad y_1 = \frac{2}{\sqrt{5}}, \quad x_2 = \frac{-2}{\sqrt{5}}, \quad y_2 = \frac{1}{\sqrt{5}}, \tag{11.8}$$

FEUE WHEP
10392

and (11.7) becomes

$$(11.9)\qquad \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}.$$

Representing these matrices by letters we can write

$$\mathrm{MC} = \mathrm{CD}, \qquad \text{where}$$

$$(11.10)\qquad \mathrm{M} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}, \quad \mathrm{C} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}, \quad \mathrm{D} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}.$$

If, as here, the determinant of C is not zero, then C has an inverse C^{-1}; let us multiply (11.10) by C^{-1} and remember that $\mathrm{C}^{-1}\mathrm{C}$ is the unit matrix; then $\mathrm{C}^{-1}\mathrm{MC} = \mathrm{C}^{-1}\mathrm{CD} = \mathrm{D}$.

$$(11.11)\qquad \mathrm{C}^{-1}\mathrm{MC} = \mathrm{D}.$$

The matrix D has elements different from zero only down the main diagonal; it is called a *diagonal matrix*. The matrix D is called *similar* to M, and when we obtain D given M, we say that we have *diagonalized* M *by a similarity transformation.*

We shall see shortly that this amounts physically to a simplification of the problem by a better choice of variables. For example, in the problem of the membrane, it is simpler to describe the deformation if we use axes along the eigenvectors. Later we shall see more examples of the use of the diagonalization process.

Observe that it is easy to find D; we need only solve the characteristic equation of M. Then D is a matrix with these characteristic values down the main diagonal and zeros elsewhere. We can also find C (with more work), but for many purposes only D is needed.

Note that the order of the eigenvalues down the main diagonal of D is arbitrary; for example we could write (11.6) as

$$(11.12)\qquad \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x_2 & x_1 \\ y_2 & y_1 \end{pmatrix} = \begin{pmatrix} x_2 & x_1 \\ y_2 & y_1 \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$$

instead of (11.7). Then (11.11) still holds, with a different C, of course, and with

$$\mathrm{D} = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$$

instead of as in (11.10) (Problem 1).

FÊUE WHEP

Meaning of C **and** D To see more clearly the meaning of (11.11) let us find what the matrices C and D mean physically. We consider two sets of axes (x, y) and (x', y') with (x', y') rotated through θ from (x, y) (Figure 11.2). The (x, y) and (x', y') coordinates of *one* point (or components of one vector $\mathbf{r} = \mathbf{r}'$) relative to the two systems are related by (7.13). Solving (7.13) for x and y, we have

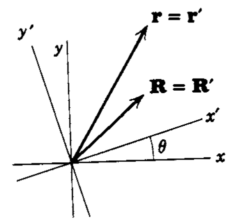


Figure 11.2

$$x = x'\cos\theta - y'\sin\theta,\qquad (11.13)$$
$$y = x'\sin\theta + y'\cos\theta,$$

or in matrix notation

$$\text{(11.14)}\qquad \mathrm{r} = \mathrm{Cr}' \qquad \text{where} \qquad \mathrm{C} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

This equation is true for *any* single vector with components given in the two systems. Suppose we have another vector $\mathbf{R} = \mathbf{R}'$ (Figure 11.2) with components X, Y and X', Y'; these components are related by

$$\text{(11.15)}\qquad \mathrm{R} = \mathrm{CR}'.$$

Now let M be a matrix which describes a deformation of the plane in the (x, y) system. Then the equation

$$\text{(11.16)}\qquad \mathrm{R} = \mathrm{Mr}$$

says that the vector $\mathbf{r}$ becomes the vector $\mathbf{R}$ after the deformation, both vectors given relative to the (x, y) axes. Let us ask how we can describe the deformation in the (x', y') system, that is, what matrix carries $\mathbf{r}'$ into $\mathbf{R}'$? We substitute (11.14) and (11.15) into (11.16) and find $\mathrm{CR}' = \mathrm{MCr}'$ or

$$\text{(11.17)}\qquad \mathrm{R}' = \mathrm{C}^{-1}\mathrm{MCr}'.$$

Thus the answer to our question is that

> $\mathrm{D} = \mathrm{C}^{-1}\mathrm{MC}$ is the matrix which describes in the (x', y') system the same deformation that M describes in the (x, y) system.

Next we want to show that if the matrix C is chosen to make $\mathrm{D} = \mathrm{C}^{-1}\mathrm{MC}$ a diagonal matrix, then the new axes (x', y') are along the directions of the eigenvectors of M. Recall from (11.10) that the columns of C are the components of the unit eigenvectors. If the eigenvectors are perpendicular, as they are in our example (see Problem 2) then the new axes (x', y') along the eigenvector directions are a set of perpendicular axes rotated

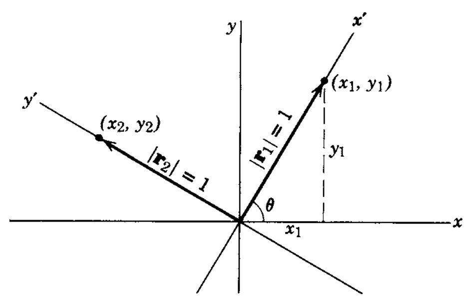


Figure 11.3

FEUE WHEP
10394

from axes (x, y) by some angle θ (Figure 11.3). The unit eigenvectors $\mathbf{r}_1$ and $\mathbf{r}_2$ are shown in Figure 11.3; from the figure we find

$$\begin{aligned} x_1 &= |\mathbf{r}_1|\cos\theta = \cos\theta, \quad x_2 = -|\mathbf{r}_2|\sin\theta = -\sin\theta \\ y_1 &= |\mathbf{r}_1|\sin\theta = \sin\theta, \quad y_2 = |\mathbf{r}_2|\cos\theta = \cos\theta; \\ \mathrm{C} &= \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \end{aligned} \tag{11.18}$$

Thus, the matrix C which diagonalizes M is the rotation matrix C in (11.14) when the (x', y') axes are along the directions of the eigenvectors of M.

Relative to these new axes, the diagonal matrix D describes the deformation. For our example we have

$$\begin{aligned} &\mathrm{R}' = \mathrm{Dr}' \qquad \text{or} \qquad \begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} \qquad \text{or} \\ &X' = x', \qquad Y' = 6y'. \end{aligned} \tag{11.19}$$

In words, (11.19) says that [in the (x', y') system] each point (x', y') has its x' coordinate unchanged by the deformation and its y' coordinate multiplied by 6, that is, the deformation is simply a stretch in the y' direction. This is a simpler description of the deformation and clearer physically than the description given by (11.1).

You can see now why the order of eigenvalues down the main diagonal in D is arbitrary and why (11.12) is just as satisfactory as (11.7). The new axes (x', y') are along the eigenvectors, but it is unimportant which eigenvector we call x' and which we call y'. In doing a problem we simply select a D with the eigenvalues of M in some (arbitrary) order down the main diagonal. Our choice of D then determines which eigenvector direction is called the x' axis and which is called y'.

It was unnecessary in the above discussion to have the x' and y' axes perpendicular, although this is the most useful case. If $\mathrm{r} = \mathrm{Cr}'$ but C is just any (nonsingular) matrix [not necessarily the orthogonal rotation matrix as in (11.14)], then (11.17) still follows. That is, $\mathrm{C}^{-1}\mathrm{MC}$ describes the deformation using (x', y') axes. But if C is not an orthogonal matrix, then the (x', y') axes are not perpendicular (Figure 11.4) and $x^2 + y^2 \neq {x'}^2 + {y'}^2$, that is, the transformation is not a rotation of axes. Recall that C is the matrix of unit eigenvectors; if these are perpendicular, then C is an orthogonal matrix (Problem 6). It can be shown that this will be the case if and only if the matrix M is symmetric. [See equation (11.27) and the discussion just before it. Also see Problems 33 to 35, and Problem 15.25.]

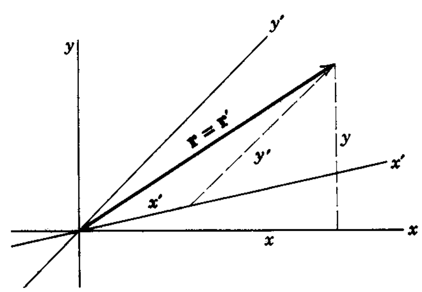


Figure 11.4

Degeneracy For a symmetric matrix, we have seen that the eigenvectors corresponding to different eigenvalues are orthogonal. If two (or more) eigenvalues are the same, then that eigenvalue is called *degenerate*. Degeneracy means that two (or more) independent eigenvectors correspond to the same eigenvalue.

10395

Example 1. Consider the following matrix:

$$M = \begin{pmatrix} 1 & -4 & 2 \\ -4 & 1 & -2 \\ 2 & -2 & -2 \end{pmatrix}. \tag{11.20}$$

The eigenvalues of M are $\lambda = 6, -3, -3$, and the eigenvector corresponding to $\lambda = 6$ is $(2, -2, 1)$ (Problem 36). For $\lambda = -3$, the eigenvector condition is $2x - 2y + z = 0$. This is a plane orthogonal to the $\lambda = 6$ eigenvector, and any vector in this plane is an eigenvector corresponding to $\lambda = -3$. That is, the $\lambda = -3$ eigenspace is a plane. It is convenient to choose two orthogonal eigenvectors as basis vectors in this $\lambda = -3$ eigenplane, for example $(1, 1, 0)$ and $(-1, 1, 4)$. (See Problem 36.)

You may ask how you find these orthogonal eigenvectors except by inspection. Recall that the cross product of two vectors is perpendicular to both of them. Thus in the present case we could pick one vector in the $\lambda = -3$ eigenplane and then take its cross product with the $\lambda = 6$ eigenvector. This gives a second vector in the $\lambda = -3$ eigenplane, perpendicular to the first one we picked. However, this only works in three dimensions; if we are dealing with spaces of higher dimension (see Section 10), then we need another method. Suppose we first write down just any two (different) vectors in the eigenplane not trying to make them orthogonal. Then we can use the Gram-Schmidt method (see Section 10) to find an orthogonal set. For example, in the problem above, suppose you had thought of (or your computer had given you) the vectors $\mathbf{A} = (1, 1, 0)$ and $\mathbf{B} = (-1, 0, 2)$ which are vectors in the $\lambda = -3$ eigenplane but not orthogonal to each other. Following the Gram-Schmidt method, we find

$$\mathbf{A} = (1, 1, 0), \qquad \mathbf{e} = \mathbf{A}/A = (1, 1, 0)/\sqrt{2},$$

$$\mathbf{B} - (\mathbf{e}\cdot\mathbf{B})\mathbf{e} = (-1, 0, 2) - \frac{-1}{2}(1, 1, 0) = \left(\frac{-1}{2}, \frac{1}{2}, 2\right),$$

or $(-1, 1, 4)$ as we had above. For a degenerate subspace of dimension $m > 2$, we just need to write down m linearly independent eigenvectors, and then find an orthogonal set by the Gram-Schmidt method.

Diagonalizing Hermitian Matrices We have seen how to diagonalize symmetric matrices by orthogonal similarity transformations. The complex analogue of a symmetric matrix ($S^T = S$) is a Hermitian matrix ($H^\dagger = H$) and the complex analogue of an orthogonal matrix ($O^T = O^{-1}$) is a unitary matrix ($U^\dagger = U^{-1}$). So let's discuss diagonalizing Hermitian matrices by unitary similarity transformations. This is of great importance in quantum mechanics.

Although Hermitian matrices may have complex off-diagonal elements, the eigenvalues of a Hermitian matrix are always real. Let's prove this. (Refer to Section 9 for definitions and theorems as needed.) Let H be a Hermitian matrix, and let r be the column matrix of a non-zero eigenvector of H corresponding to the eigenvalue λ. Then the eigenvector condition is $Hr = \lambda r$. We want to take the transpose conjugate (dagger) of this equation. Using the complex conjugate of equation (9.10), we get $(Hr)^\dagger = r^\dagger H^\dagger = r^\dagger H$ since $H^\dagger = H$ for a Hermitian matrix. The transpose conjugate of λr is $\lambda^* r^\dagger$ (since λ is a number, we just need to take its complex conjugate). Now we have the two equations

$$Hr = \lambda r \qquad \text{and} \qquad r^\dagger H = \lambda^* r^\dagger. \tag{11.21}$$

FEUE WHEP
10396

Multiply the first equation in (11.21) on the left [see discussion following (10.9)] by the row matrix $\mathrm{r}^\dagger$ and the second equation on the right by the column matrix r to get

$$\text{(11.22)} \qquad \mathrm{r}^\dagger \mathrm{H}\mathrm{r} = \lambda \mathrm{r}^\dagger \mathrm{r} \quad \text{and} \quad \mathrm{r}^\dagger \mathrm{H}\mathrm{r} = \lambda^* \mathrm{r}^\dagger \mathrm{r}.$$

Subtracting the two equations we find $(\lambda - \lambda^*)\mathrm{r}^\dagger\mathrm{r} = 0$. Since we assumed $\mathrm{r} \neq 0$, we have $\lambda^* = \lambda$, that is, λ is real.

We can also show that for a Hermitian matrix the eigenvectors corresponding to two different eigenvalues are orthogonal. Start with the two eigenvector conditions,

$$\text{(11.23)} \qquad \mathrm{H}\mathrm{r}_1 = \lambda_1 \mathrm{r}_1 \quad \text{and} \quad \mathrm{H}\mathrm{r}_2 = \lambda_2 \mathrm{r}_2.$$

From these we can show (Problem 37)

$$\text{(11.24)} \qquad \mathrm{r}_1^\dagger \mathrm{H}\mathrm{r}_2 = \lambda_1 \mathrm{r}_1^\dagger \mathrm{r}_2 = \lambda_2 \mathrm{r}_1^\dagger \mathrm{r}_2, \quad \text{or} \quad (\lambda_1 - \lambda_2)\mathrm{r}_1^\dagger \mathrm{r}_2 = 0.$$

Thus if $\lambda_1 \neq \lambda_2$, then the inner product of r_1 and r_2 is zero, that is, they are orthogonal [see (10.8)].

We can also prove that if a matrix M has real eigenvalues and can be diagonalized by a unitary similarity transformation, then it is Hermitian. In symbols, we write $\mathrm{U}^{-1}\mathrm{M}\mathrm{U} = \mathrm{D}$, and find the transpose conjugate of this equation to get (Problem 38)

$$\text{(11.25)} \qquad (\mathrm{U}^{-1}\mathrm{M}\mathrm{U})^\dagger = \mathrm{U}^{-1}\mathrm{M}^\dagger\mathrm{U} = \mathrm{D}^\dagger = \mathrm{D}.$$

Thus $\mathrm{U}^{-1}\mathrm{M}\mathrm{U} = \mathrm{D} = \mathrm{U}^{-1}\mathrm{M}^\dagger\mathrm{U}$, so $\mathrm{M} = \mathrm{M}^\dagger$, which says that M is Hermitian. So we have proved that

(11.26) A matrix has real eigenvalues and can be diagonalized by a unitary similarity transformation if and only if it is Hermitian.

Since a real Hermitian matrix is a symmetric matrix and a real unitary matrix is an orthogonal matrix, the corresponding statement for symmetric matrices is (Problem 39).

(11.27) A matrix has real eigenvalues and can be diagonalized by an orthogonal similarity transformation if and only if it is symmetric.

Recall from (9.2) and Problem 9.22 that normal matrices include symmetric, Hermitian, orthogonal, and unitary matrices (as well as some others). It may be useful to know the following general theorem which we state without proof [see, for example, Am. J. Phys. **52**, 513–515 (1984)].

(11.28) A matrix can be diagonalized by a unitary similarity transformation if and only if it is normal.

FÊUE WHEP

Example 2. To illustrate diagonalizing a Hermitian matrix by a unitary similarity transformation, we consider the matrix

$$\mathrm{H} = \begin{pmatrix} 2 & 3-i \\ 3+i & -1 \end{pmatrix}. \tag{11.29}$$

(Verify that H is Hermitian.) We follow the same routine we used to find the eigenvalues and eigenvectors of a symmetric matrix. The eigenvalues are given by

$$(2-\lambda)(-1-\lambda) - (3+i)(3-i) = 0,$$
$$\lambda^2 - \lambda - 12 = 0, \quad \lambda = -3, 4.$$

For $\lambda = -3$, an eigenvector satisfies the equations

$$\begin{pmatrix} 5 & 3-i \\ 3+i & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0, \quad \text{or}$$
$$5x + (3-i)y = 0, \quad (3+i)x + 2y = 0.$$

These equations are satisfied by $x = 2$, $y = (-3-i)$. A choice for the unit eigenvector is $(2, -3-i)/\sqrt{14}$. For $\lambda = 4$, we find similarly the equations

$$-2x + (3-i)y = 0, \quad (3+i)x - 5y = 0,$$

which are satisfied by $y = 2$, $x = 3-i$, so a unit eigenvector is $(3-i, 2)/\sqrt{14}$. We can verify that the two eigenvectors are orthogonal (as we proved above that they must be) by finding that their inner product [see (10.8)] is $(2, -3-i)^* \cdot (3-i, 2) = 2(3-i) + 2(-3+i) = 0$. As in (11.10) we write the unit eigenvectors as the columns of a matrix U which diagonalizes H by a similarity transformation.

$$\mathrm{U} = \frac{1}{\sqrt{14}} \begin{pmatrix} 2 & 3-i \\ -3-i & 2 \end{pmatrix}, \quad \mathrm{U}^\dagger = \frac{1}{\sqrt{14}} \begin{pmatrix} 2 & -3+i \\ 3+i & 2 \end{pmatrix}$$

You can easily verify that $\mathrm{U}^\dagger\mathrm{U}$ = the unit matrix, so $\mathrm{U}^{-1} = \mathrm{U}^\dagger$. Then (Problem 40)

$$\mathrm{U}^{-1}\mathrm{H}\mathrm{U} = \mathrm{U}^\dagger\mathrm{H}\mathrm{U} = \begin{pmatrix} -3 & 0 \\ 0 & 4 \end{pmatrix}, \tag{11.30}$$

that is, H is diagonalized by a unitary similarity transformation.

Orthogonal Transformations in 3 Dimensions In Section 7, we considered the active rotation and/or reflection of vectors $\mathbf{r}$ which was produced by a given 3 by 3 orthogonal matrix. Study Equations (7.18) and (7.19) carefully to see that, acting on a column vector $\mathbf{r}$, they rotate the vector by angle θ around the z axis and/or reflect it through the (x, y) plane. We would now like to see how to find the effect of more complicated orthogonal matrices. We can do this by using an orthogonal similarity transformation to write a given orthogonal matrix relative to a new coordinate system in which the rotation axis is the z axis, and/or the (x, y) plane is the reflecting plane (in vector space language, this is a change of basis). Then a comparison with (7.18) or (7.19) gives the rotation angle. Recall how we construct a C matrix so that $\mathrm{C}^{-1}\mathrm{MC}$ describes the same transformation relative to a new set of axes that M described relative to the original axes: The columns of the C matrix are the components of unit vectors along the new axes [see (11.18) and Figure 11.3].

10398

▶ **Example 3.** Consider the following matrices.

$$\text{(11.31)} \qquad A = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix}, \quad B = \frac{1}{3}\begin{pmatrix} -2 & -1 & -2 \\ 2 & -2 & -1 \\ 1 & 2 & -2 \end{pmatrix}$$

You can verify that A and B are both orthogonal, and that $\det A = 1$, $\det B = -1$ (Problem 45). Thus A is a rotation matrix while B involves a reflection (and perhaps also a rotation). For A, a vector along the rotation axis is not affected by the transformation so we find the rotation axis by solving the equation $Ar = r$. We did this in Section 7, but now you should recognize this as an eigenvector equation. We want the eigenvector corresponding to the eigenvalue 1. By hand or by computer (Problem 45) we find that the eigenvector of A corresponding to $\lambda = 1$ is $(1, 0, 1)$ or $\mathbf{i} + \mathbf{k}$; this is the rotation axis. We want the new z axis to lie along this direction, so we take the elements of the third column of matrix C to be the components of the unit vector $\mathbf{u} = (1, 0, 1)/\sqrt{2}$. For the first column (new x axis) we choose a unit vector perpendicular to the rotation axis, say $\mathbf{v} = (1, 0, -1)/\sqrt{2}$, and for the second column (new y axis), we use the cross product $\mathbf{u} \times \mathbf{v} = (0, 1, 0)$ (so that the new axes form a right-handed orthogonal triad). This gives (Problem 45)

$$\text{(11.32)} \qquad C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad C^{-1}AC = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Comparing this result with (7.18), we see that $\cos\theta = 0$ and $\sin\theta = -1$, so the rotation is $-90°$ around the axis $\mathbf{i} + \mathbf{k}$ (or, if you prefer, $+90°$ around $-\mathbf{i} - \mathbf{k}$).

▶ **Example 4.** For the matrix B, a vector perpendicular to the reflection plane is reversed in direction by the reflection. Thus we want to solve the equation $Br = -r$, that is, to find the eigenvector corresponding to $\lambda = -1$. You can verify (Problem 45) that this is the vector $(1, -1, 1)$ or $\mathbf{i} - \mathbf{j} + \mathbf{k}$. The reflection is through the plane $x - y + z = 0$, and the rotation (if any) is about the vector $\mathbf{i} - \mathbf{j} + \mathbf{k}$. As we did for matrix A, we construct a matrix C from this vector and two perpendicular vectors, to get (Problem 45)

$$\text{(11.33)} \qquad C = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{-2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}, \qquad C^{-1}BC = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Compare this with (7.19) to get $\cos\theta = -\frac{1}{2}$, $\sin\theta = \frac{\sqrt{3}}{2}$, so matrix B produces a rotation of $120°$ around $\mathbf{i} - \mathbf{j} + \mathbf{k}$ and a reflection through the plane $x - y + z = 0$.

You may have discovered that matrices A and B have two complex eigenvalues (see Problem 46). The corresponding eigenvectors are also complex, and we didn't use them because this would take us into complex vector space (see Section 10, and Problem 47) and our rotation and reflection problems are in ordinary real 3-dimensional space. (Note also that we did not diagonalize A and B, but just

FEUE WHEP

used similarity transformations to display them relative to rotated axes.) However, when all the eigenvalues of an orthogonal matrix are real (see Problem 48), then this process produces a diagonalized matrix with the eigenvalues down the main diagonal.

Example 5. Consider the matrix

$$\mathrm{F} = \frac{1}{7}\begin{pmatrix} 2 & 6 & 3 \\ 6 & -3 & 2 \\ 3 & 2 & -6 \end{pmatrix}. \tag{11.34}$$

You can verify (Problem 49) that $\det \mathrm{F} = 1$, that the rotation axis (eigenvector corresponding to the eigenvalue $\lambda = 1$) is $3\mathbf{i} + 2\mathbf{j} + \mathbf{k}$, and that the other two eigenvalues are $-1, -1$. Then the diagonalized F (relative to axes with the new z axis along the rotation axis) is

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{11.35}$$

Comparing this with equation (7.18), we see that $\cos\theta = -1$, $\sin\theta = 0$, so F produces a rotation of $180°$ about $3\mathbf{i} + 2\mathbf{j} + \mathbf{k}$.

An even easier way to find the rotation angle in this problem is to use the trace of F (Problem 50). From (7.18) and (11.34) we have $2\cos\theta + 1 = -1$. Thus $\cos\theta = -1$, $\theta = 180°$ as before. This method gives $\cos\theta$ for any rotation or reflection matrix, but unless $\cos\theta = \pm 1$, we also need more information (say the value of $\sin\theta$) to determine whether θ is positive or negative.

Powers and Functions of Matrices In Section 6 we found functions of some matrices A for which it was easy to find the powers because they repeated periodically [see equations (6.15) to (6.17)]. When this doesn't happen, it isn't so easy to find powers directly (Problem 58). But it is easy to find powers of a diagonal matrix, and you can also show that (Problem 57)

$$\mathrm{M}^n = \mathrm{C}\mathrm{D}^n\mathrm{C}^{-1}, \quad \text{where} \quad \mathrm{C}^{-1}\mathrm{M}\mathrm{C} = \mathrm{D}, \quad \mathrm{D}\ \text{diagonal}. \tag{11.36}$$

This result is useful not just for evaluating powers and functions of numerical matrices but also for proving theorems (Problem 60).

Example 6. We can show that if, as above, $\mathrm{C}^{-1}\mathrm{M}\mathrm{C} = \mathrm{D}$, then

$$\det e^{\mathrm{M}} = e^{\mathrm{Tr}(\mathrm{M})}. \tag{11.37}$$

As in (6.17) we define e^{M} by its power series. For each term of the series $\mathrm{M}^n = \mathrm{C}\mathrm{D}^n\mathrm{C}^{-1}$ by (11.36), so $e^{\mathrm{M}} = \mathrm{C}e^{\mathrm{D}}\mathrm{C}^{-1}$. By (6.6), the determinant of a product = the product of the determinants, and $\det \mathrm{C}\mathrm{C}^{-1} = 1$, so we have $\det e^{\mathrm{M}} = \det e^{\mathrm{D}}$. Now the matrix e^{D} is diagonal and the diagonal elements are e^{λ_i} where λ_i are the eigenvalues of M. Thus $\det e^{\mathrm{D}} = e^{\lambda_1}e^{\lambda_2}e^{\lambda_3}\cdots = e^{\mathrm{Tr}\,\mathrm{D}}$. But by (9.13), $\mathrm{Tr}\,\mathrm{D} = \mathrm{Tr}(\mathrm{C}\mathrm{C}^{-1}\mathrm{M}) = \mathrm{Tr}\,\mathrm{M}$, so we have (11.37).

FEUE WHEP

Simultaneous Diagonalization Can we diagonalize two (or more) matrices using the same similarity transformation? Sometimes we can, namely if, and only if, they commute. Let's see why this is true. Recall that the diagonalizing C matrix has columns which are mutually orthogonal unit eigenvectors of the matrix being diagonalized. Suppose we can find the same set of eigenvectors for two matrices F and G; then the same C will diagonalize both. So the problem amounts to showing how to find a common set of eigenvectors for F and G if they commute.

▶ **Example 7.** Let's start by diagonalizing F. Suppose r (a column matrix) is the eigenvector corresponding to the eigenvalue λ, that is, $\mathrm{Fr} = \lambda\mathrm{r}$. Multiply this on the left by G and use $\mathrm{GF} = \mathrm{FG}$ (matrices commute) to get

$$\mathrm{GFr} = \lambda\mathrm{Gr}, \qquad \text{or} \qquad \mathrm{F(Gr)} = \lambda(\mathrm{Gr}). \tag{11.38}$$

This says that Gr is an eigenvector of F corresponding to the eigenvalue λ. If λ is not degenerate (that is if there is just one eigenvector corresponding to λ) then Gr must be the same vector as r (except maybe for length), that is, Gr is a multiple of r, or $\mathrm{Gr} = \lambda'\mathrm{r}$. This is the eigenvector equation for G; it says that r is an eigenvector of G. If all eigenvalues of F are non-degenerate, then F and G have the same set of eigenvectors, and so can be diagonalized by the same C matrix.

▶ **Example 8.** Now suppose that there are two (or more) linearly independent eigenvectors corresponding to the eigenvalue λ of F. Then every vector in the degenerate eigenspace corresponding to λ is an eigenvector of matrix F (see discussion of degeneracy above). Next consider matrix G. Corresponding to all non-degenerate F eigenvalues we already have the same set of eigenvectors for G as for F. So we just have to find the eigenvectors of G in the degenerate eigenspace of F. Since all vectors in this subspace are eigenvectors of F, we are free to choose ones which are eigenvectors of G. Thus we now have the same set of eigenvectors for both matrices, and so we can construct a C matrix which will diagonalize both F and G. For the converse, see Problem 62.

▶ PROBLEMS, SECTION 11

1. Verify (11.7). Also verify (11.12) and find the corresponding different C in (11.11). *Hint:* To find C, start with (11.12) instead of (11.7) and follow through the method of getting (11.10) from (11.7).
2. Verify that the two eigenvectors in (11.8) are perpendicular, and that C in (11.10) satisfies the condition (7.9) for an orthogonal matrix.
3. **(a)** If C is orthogonal and M is symmetric, show that $\mathrm{C}^{-1}\mathrm{MC}$ is symmetric.
   (b) If C is orthogonal and M antisymmetric, show that $\mathrm{C}^{-1}\mathrm{MC}$ is antisymmetric.
4. Find the inverse of the rotation matrix in (7.13); you should get C in (11.14). Replace θ by $-\theta$ in (7.13) to see that the matrix C corresponds to a rotation through $-\theta$.
5. Show that the C matrix in (11.10) does represent a rotation by finding the rotation angle. Write equations (7.13) and (11.13) for this rotation.
6. Show that if C is a matrix whose columns are the components (x_1, y_1) and (x_2, y_2) of two perpendicular vectors each of unit length, then C is an orthogonal matrix. *Hint:* Find $\mathrm{C}^{\mathrm{T}}\mathrm{C}$.

10401

7. Generalize Problem 6 to three dimensions; to n dimensions.

8. Show that under the transformation (11.1), all points (x, y) on a given straight line through the origin go into points (X, Y) on another straight line through the origin. *Hint:* Solve (11.1) for x and y in terms of X and Y and substitute into the equation $y = mx$ to get an equation $Y = kX$, where k is a constant. *Further hint:* If $\mathrm{R} = \mathrm{Mr}$, then $\mathrm{r} = \mathrm{M}^{-1}\mathrm{R}$.

9. Show that $\det(\mathrm{C}^{-1}\mathrm{MC}) = \det \mathrm{M}$. *Hints:* See (6.6). What is the product of $\det(\mathrm{C}^{-1})$ and $\det \mathrm{C}$? Thus show that the product of the eigenvalues of M is equal to $\det \mathrm{M}$.

10. Show that $\mathrm{Tr}(\mathrm{C}^{-1}\mathrm{MC}) = \mathrm{Tr}\,\mathrm{M}$. *Hint:* See (9.13). Thus show that the sum of the eigenvalues of M is equal to $\mathrm{Tr}\,\mathrm{M}$.

11. Find the inverse of the transformation $x' = 2x - 3y$, $y' = x + y$, that is, find x, y in terms of x', y'. (*Hint:* Use matrices.) Is the transformation orthogonal?

Find the eigenvalues and eigenvectors of the following matrices. Do some problems by hand to be sure you understand what the process means. Then check your results by computer.

12. $\begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}$ **13.** $\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}$ **14.** $\begin{pmatrix} 3 & -2 \\ -2 & 0 \end{pmatrix}$

15. $\begin{pmatrix} 2 & 3 & 0 \\ 3 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ **16.** $\begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & -1 \end{pmatrix}$ **17.** $\begin{pmatrix} 5 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 5 \end{pmatrix}$

18. $\begin{pmatrix} -1 & 1 & 3 \\ 1 & 2 & 0 \\ 3 & 0 & 2 \end{pmatrix}$ **19.** $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 0 \\ 2 & 0 & 3 \end{pmatrix}$ **20.** $\begin{pmatrix} -1 & 2 & 1 \\ 2 & 3 & 0 \\ 1 & 0 & 3 \end{pmatrix}$

21. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$ **22.** $\begin{pmatrix} -3 & 2 & 2 \\ 2 & 1 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ **23.** $\begin{pmatrix} 13 & 4 & -2 \\ 4 & 13 & -2 \\ -2 & -2 & 10 \end{pmatrix}$

24. $\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$ **25.** $\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & -1 \end{pmatrix}$ **26.** $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$

Let each of the following matrices M describe a deformation of the (x, y) plane. For each given M find: the eigenvalues and eigenvectors of the transformation, the matrix C which diagonalizes M and specifies the rotation to new axes (x', y') along the eigenvectors, and the matrix D which gives the deformation relative to the new axes. Describe the deformation relative to the new axes.

27. $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ **28.** $\begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$ **29.** $\begin{pmatrix} 3 & 4 \\ 4 & 9 \end{pmatrix}$

30. $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ **31.** $\begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$ **32.** $\begin{pmatrix} 6 & -2 \\ -2 & 3 \end{pmatrix}$

33. Find the eigenvalues and eigenvectors of the real symmetric matrix

$$\mathrm{M} = \begin{pmatrix} A & H \\ H & B \end{pmatrix}.$$

Show that the eigenvalues are real and the eigenvectors are perpendicular.

FEUE WHEP
10402

34. By multiplying out $M = CDC^{-1}$ where C is the rotation matrix (11.14) and D is the diagonal matrix

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

show that if M can be diagonalized by a rotation, then M is symmetric.

35. The characteristic equation for a second-order matrix M is a quadratic equation. We have considered in detail the case in which M is a real symmetric matrix and the roots of the characteristic equation (eigenvalues) are real, positive, and unequal. Discuss some other possibilities as follows:

(a) M real and symmetric, eigenvalues real, one positive and one negative. Show that the plane is reflected in one of the eigenvector lines (as well as stretched or shrunk). Consider as a simple special case

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(b) M real and symmetric, eigenvalues equal (and therefore real). Show that M must be a multiple of the unit matrix. Thus show that the deformation consists of dilation or shrinkage in the radial direction (the same in all directions) with no rotation (and reflection in the origin if the root is negative).

(c) M real, *not* symmetric, eigenvalues real and not equal. Show that in this case the eigenvectors are not orthogonal. *Hint:* Find their dot product.

(d) M real, *not* symmetric, eigenvalues complex. Show that all vectors are rotated, that is, there are no (real) eigenvectors which are unchanged in direction by the transformation. Consider the characteristic equation of a rotation matrix as a special case.

36. Verify the eigenvalues and eigenvectors of matrix M in (11.20). Find some other pairs of orthogonal eigenvectors in the $\lambda = -3$ eigenplane.

37. Starting with (11.23), obtain (11.24). *Hints:* Take the transpose conjugate (dagger) of the first equation in (11.23), (remember that H is Hermitian and the λ's are real) and multiply on the right by r_2. Multiply the second equation in (11.23) on the left by $r_1^\dagger$.

38. Verify equation (11.25). *Hint:* Remember from Section 9 that the transpose conjugate (dagger) of a product of matrices is the product of the transpose conjugates in reverse order and that $U^\dagger = U^{-1}$. Also remember that we have assumed real eigenvalues, so D is a real diagonal matrix.

39. Write out the detailed proof of (11.27). *Hint:* Follow the proof of (11.26) in equations (11.21) to (11.25), replacing the Hermitian matrix H by a symmetric matrix M which is real. However, don't assume that the eigenvalues λ are real until you prove it.

40. Verify the details as indicated in diagonalizing H in (11.29).

Verify that each of the following matrices is Hermitian. Find its eigenvalues and eigenvectors, write a unitary matrix U which diagonalizes H by a similarity transformation, and show that $U^{-1}HU$ is the diagonal matrix of eigenvalues.

41. $\begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$

42. $\begin{pmatrix} 3 & 1-i \\ 1+i & 2 \end{pmatrix}$

43. $\begin{pmatrix} 1 & 2i \\ -2i & -2 \end{pmatrix}$

44. $\begin{pmatrix} -2 & 3+4i \\ 3-4i & -2 \end{pmatrix}$

FEUE WHEP

45. Verify the details in the discussion of the matrices in (11.31).

46. We have seen that an orthogonal matrix with determinant 1 has at least one eigenvalue $= 1$, and an orthogonal matrix with determinant $= -1$ has at least one eigenvalue $= -1$. Show that the other two eigenvalues in both cases are $e^{i\theta}$, $e^{-i\theta}$, which, of course, includes the real values 1 (when $\theta = 0$), and -1 (when $\theta = \pi$). *Hint:* See Problem 9, and remember that rotations and reflections do not change the length of vectors so eigenvalues must have absolute value $= 1$.

47. Find a unitary matrix U which diagonalizes A in (11.31) and verify that $\mathrm{U}^{-1}\mathrm{AU}$ is diagonal with the eigenvalues down the main diagonal.

48. Show that an orthogonal matrix M with all real eigenvalues is symmetric. *Hints*: Method 1. When the eigenvalues are real, so are the eigenvectors, and the unitary matrix which diagonalizes M is orthogonal. Use (11.27). Method 2. From Problem 46, note that the only real eigenvalues of an orthogonal M are ± 1. Thus show that $\mathrm{M} = \mathrm{M}^{-1}$. Remember that M is orthogonal to show that $\mathrm{M} = \mathrm{M}^{\mathrm{T}}$.

49. Verify the results for F in the discussion of (11.34).

50. Show that the trace of a rotation matrix equals $2\cos\theta + 1$ where θ is the rotation angle, and the trace of a reflection matrix equals $2\cos\theta - 1$. *Hint*: See equations (7.18) and (7.19), and Problem 10.

Show that each of the following matrices is orthogonal and find the rotation and/or reflection it produces as an operator acting on vectors. If a rotation, find the axis and angle; if a reflection, find the reflecting plane and the rotation, if any, about the normal to that plane.

51. $\dfrac{1}{11}\begin{pmatrix} 2 & 6 & 9 \\ 6 & 7 & -6 \\ 9 & -6 & 2 \end{pmatrix}$

52. $\dfrac{1}{2}\begin{pmatrix} -1 & -1 & \sqrt{2} \\ 1 & 1 & \sqrt{2} \\ \sqrt{2} & -\sqrt{2} & 0 \end{pmatrix}$

53. $\dfrac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}$

54. $\dfrac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & -1 \\ \sqrt{2} & 0 & \sqrt{2} \\ 1 & -\sqrt{2} & -1 \end{pmatrix}$

55. $\dfrac{1}{9}\begin{pmatrix} -1 & 8 & 4 \\ -4 & -4 & 7 \\ -8 & 1 & -4 \end{pmatrix}$

56. $\dfrac{1}{2\sqrt{2}}\begin{pmatrix} 2 & \sqrt{2} & \sqrt{2} \\ -\sqrt{2} & 1+\sqrt{2} & 1-\sqrt{2} \\ -\sqrt{2} & 1-\sqrt{2} & 1+\sqrt{2} \end{pmatrix}$

57. Show that if D is a diagonal matrix, then D^n is the diagonal matrix with elements equal to the n^{th} power of the elements of D. Also show that if $\mathrm{D} = \mathrm{C}^{-1}\mathrm{MC}$, then $\mathrm{D}^n = \mathrm{C}^{-1}\mathrm{M}^n\mathrm{C}$, so $\mathrm{M}^n = \mathrm{CD}^n\mathrm{C}^{-1}$. *Hint*: For $n = 2$, $(\mathrm{C}^{-1}\mathrm{MC})^2 = \mathrm{C}^{-1}\mathrm{MCC}^{-1}\mathrm{MC}$; what is CC^{-1}?

58. Note in Section 6 [see (6.15)] that, for the given matrix A, we found $\mathrm{A}^2 = -\mathrm{I}$, so it was easy to find all the powers of A. It is not usually this easy to find high powers of a matrix directly. Try it for the square matrix M in equation (11.1). Then use the method outlined in Problem 57 to find M^4, M^{10}, e^{M}.

59. Repeat the last part of Problem 58 for the matrix $\mathrm{M} = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$.

60. The Caley-Hamilton theorem states that "A matrix satisfies its own characteristic equation." Verify this theorem for the matrix M in equation (11.1). *Hint*: Substitute the matrix M for λ in the characteristic equation (11.4) and verify that you have a correct matrix equation. *Further hint*: Don't do all the arithmetic. Use (11.36) to write the left side of your equation as $\mathrm{C}(\mathrm{D}^2 - 7\mathrm{D} + 6)\mathrm{C}^{-1}$ and show that the parenthesis $= 0$. Remember that, by definition, the eigenvalues satisfy the characteristic equation.

FEUE WHEP
10404

61. At the end of Section 9 we proved that if H is a Hermitian matrix, then the matrix $e^{i\mathrm{H}}$ is unitary. Give another proof by writing $\mathrm{H} = \mathrm{CDC}^{-1}$, remembering that now C is unitary and the eigenvalues in D are real. Show that $e^{i\mathrm{D}}$ is unitary and that $e^{i\mathrm{H}}$ is a product of three unitary matrices. See Problem 9.17d.

62. Show that if matrices F and G can be diagonalized by the same C matrix, then they commute. *Hint*: Do diagonal matrices commute?

▶ 12. APPLICATIONS OF DIAGONALIZATION

We next consider some examples of the use of the diagonalization process. A central conic section (ellipse or hyperbola) with center at the origin has the equation

$$Ax^2 + 2Hxy + By^2 = K, \tag{12.1}$$

where A, B, H and K are constants. In matrix form this can be written

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} A & H \\ H & B \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = K \quad \text{or} \quad \begin{pmatrix} x & y \end{pmatrix} \mathrm{M} \begin{pmatrix} x \\ y \end{pmatrix} = K \tag{12.2}$$

if we call

$$\begin{pmatrix} A & H \\ H & B \end{pmatrix} = \mathrm{M}$$

(as you can verify by multiplying out the matrices). We want to choose the principal axes of the conic as our reference axes in order to write the equation in simpler form. Consider Figure 11.2; let the axes (x', y') be rotated by some angle θ from (x, y). Then the (x, y) and (x', y') coordinates of a point are related by (11.13) or (11.14):

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \mathrm{C} \begin{pmatrix} x' \\ y' \end{pmatrix}. \tag{12.3}$$

By (9.11) the transpose of (12.3) is

$$\begin{aligned} &\begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \\ &\begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} x' & y' \end{pmatrix} \mathrm{C}^{\mathrm{T}} = \begin{pmatrix} x' & y' \end{pmatrix} \mathrm{C}^{-1} \end{aligned} \tag{12.4}$$

since C is an orthogonal matrix. Substituting (12.3) and (12.4) into (12.2), we get

$$\begin{pmatrix} x' & y' \end{pmatrix} \mathrm{C}^{-1}\mathrm{MC} \begin{pmatrix} x' \\ y' \end{pmatrix} = K. \tag{12.5}$$

If C is the matrix which diagonalizes M, then (12.5) is the equation of the conic relative to its principal axes.

▶ **Example 1.** Consider the conic

$$5x^2 - 4xy + 2y^2 = 30. \tag{12.6}$$

In matrix form this can be written

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 30. \tag{12.7}$$

FEUE WHEP
10405

We have here the same matrix,

$$\mathrm{M} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix},$$

whose eigenvalues we found in Section 11. In that section we found a C such that

$$\mathrm{C}^{-1}\mathrm{MC} = \mathrm{D} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}.$$

Then the equation (12.5) of the conic relative to principal axes is

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = x'^2 + 6y'^2 = 30. \tag{12.8}$$

Observe that changing the order of 1 and 6 in D would give $6x'^2 + y'^2 = 30$ as the new equation of the ellipse instead of (12.8). This amounts simply to interchanging the x' and y' axes.

By comparing the matrix C of the unit eigenvectors in (11.10) with the rotation matrix in (11.14), we see that the rotation angle θ (Figure 11.3) from the original axes (x, y) to the principal axes (x', y') is

$$\theta = \arccos \frac{1}{\sqrt{5}}. \tag{12.9}$$

Notice that in writing the conic section equation in matrix form (12.2) and (12.7), we split the xy term evenly between the two nondiagonal elements of the matrix; this made M symmetric. Recall (end of Section 11) that M can be diagonalized by a similarity transformation $\mathrm{C}^{-1}\mathrm{MC}$ with C an orthogonal matrix (that is, by a rotation of axes) if and only if M is symmetric. We choose M symmetric (by splitting the xy term in half) to make our process work.

Although for simplicity we have been working in two dimensions, the same ideas apply to three (or more) dimensions (that is, three or more variables). As we have said (Section 10), although we can represent only three coordinates in physical space, it is very convenient to use the same geometrical terminology even though the number of variables is greater than three. Thus if we diagonalize a matrix of any order, we still use the terms eigenvalues, eigenvectors, principal axes, rotation to principal axes, etc.

Example 2. Rotate to principal axes the quadric surface

$$x^2 + 6xy - 2y^2 - 2yz + z^2 = 24.$$

In matrix form this equation is

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 24.$$

The characteristic equation of this matrix is

$$\begin{vmatrix} 1-\lambda & 3 & 0 \\ 3 & -2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = 0 = -\lambda^3 + 13\lambda - 12$$
$$= -(\lambda - 1)(\lambda + 4)(\lambda - 3).$$

10406

The characteristic values are

$$\lambda = 1, \qquad \lambda = -4, \qquad \lambda = 3.$$

Relative to the principal axes (x', y', z') the quadric surface equation becomes

$$\begin{pmatrix} x' & y' & z' \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = 24$$

or

$$x'^2 - 4y'^2 + 3z'^2 = 24.$$

From this equation we can identify the quadric surface (hyperboloid of one sheet) and sketch its size and shape using (x', y', z') axes without finding their relation to the original (x, y, z) axes. However, if we do want to know the relation between the two sets of axes, we find the C matrix in the following way. Recall from Section 11 that C is the matrix whose columns are the components of the unit eigenvectors. One of the eigenvectors can be found by substituting the eigenvalue $\lambda = 1$ into the equations

$$\begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \\ \lambda z \end{pmatrix}$$

and solving for x, y, z. Then $\mathbf{i}x + \mathbf{j}y + \mathbf{k}z$ is an eigenvector corresponding to $\lambda = 1$, and by dividing it by its magnitude we get a *unit* eigenvector (Problem 8). Repeating this process for each of the other values of λ, we get the following three unit eigenvectors:

$$\left(\frac{1}{\sqrt{10}}, 0, \frac{3}{\sqrt{10}}\right) \qquad \text{when} \quad \lambda = 1;$$

$$\left(\frac{-3}{\sqrt{35}}, \frac{5}{\sqrt{35}}, \frac{1}{\sqrt{35}}\right) \qquad \text{when} \quad \lambda = -4;$$

$$\left(\frac{-3}{\sqrt{14}}, \frac{-2}{\sqrt{14}}, \frac{1}{\sqrt{14}}\right) \qquad \text{when} \quad \lambda = 3.$$

Then the rotation matrix C is

$$\mathrm{C} = \begin{pmatrix} \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{35}} & \frac{-3}{\sqrt{14}} \\ 0 & \frac{5}{\sqrt{35}} & \frac{-2}{\sqrt{14}} \\ \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{35}} & \frac{1}{\sqrt{14}} \end{pmatrix}$$

The numbers in C are the cosines of the nine angles between the (x, y, z) and (x', y', z') axes. (Compare Figure 11.3 and the discussion of it.)

A useful physical application of this method occurs in discussing vibrations. We illustrate this with a simple problem.

FEUE WHEP

Example 3. Find the characteristic vibration frequencies for the system of masses and springs shown in Figure 12.1.

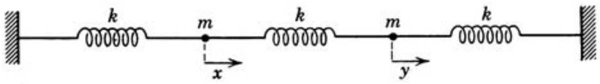


Figure 12.1

Let x and y be the coordinates of the two masses at time t relative to their equilibrium positions, as shown in Figure 12.1. We want to write the equations of motion (mass times acceleration = force) for the two masses (see Chapter 2, end of Section 16). We *can* just write the forces by inspection as we did in Chapter 2, but for more complicated problems it is useful to have a systematic method. First write the potential energy; for a spring this is $V = \frac{1}{2}ky^2$ where y is the compression or extension of the spring from its equilibrium length. Then the force exerted on a mass attached to the spring is $-ky = -dV/dy$. If V is a function of two (or more) variables, say x and y as in Figure 12.1, then the forces on the two masses are $-\partial V/\partial x$ and $-\partial V/\partial y$ (and so on for more variables). For Figure 12.1, the extension or compression of the middle spring is $x - y$ so its potential energy is $\frac{1}{2}k(x-y)^2$. For the other two springs, the potential energies are $\frac{1}{2}kx^2$ and $\frac{1}{2}ky^2$ so the total potential energy is

$$V = \frac{1}{2}kx^2 + \frac{1}{2}k(x-y)^2 + \frac{1}{2}ky^2 = k(x^2 - xy + y^2). \tag{12.10}$$

In writing the equations of motion it is convenient to use a dot to indicate a time derivative (as we often use a prime to mean an x derivative). Thus $\dot{x} = dx/dt$, $\ddot{x} = d^2x/dt^2$, etc. Then the equations of motion are

$$\begin{cases} m\ddot{x} = -\partial V/\partial x = -2kx + ky, \\ m\ddot{y} = -\partial V/\partial y = \quad kx - 2ky. \end{cases} \tag{12.11}$$

In a *normal* or *characteristic* mode of vibration, the x and y vibrations have the same frequency. As in Chapter 2, equations (16.22), we assume solutions $x = x_0e^{i\omega t}$, $y = y_0e^{i\omega t}$, with the same frequency ω for both x and y. [Or, if you prefer, we could replace $e^{i\omega t}$ by $\sin\omega t$ or $\cos\omega t$ or $\sin(\omega t + \alpha)$, etc.] Note that (for any of these solutions),

$$\ddot{x} = -\omega^2 x, \quad \text{and} \quad \ddot{y} = -\omega^2 y. \tag{12.12}$$

Substituting (12.12) into (12.11) we get (Problem 10)

$$\begin{cases} -m\omega^2 x = -2kx + ky, \\ -m\omega^2 y = \quad kx - 2ky. \end{cases} \tag{12.13}$$

In matrix form these equations are

$$\lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with} \quad \lambda = \frac{m\omega^2}{k}. \tag{12.14}$$

Note that this is an eigenvalue problem (see Section 11). To find the eigenvalues λ, we write

$$\begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = 0 \tag{12.15}$$

FEUE WHEP
10408

and solve for λ to find $\lambda = 1$ or $\lambda = 3$. Thus [by the definition of λ in (12.14)] the characteristic frequencies are

$$\omega_1 = \sqrt{\frac{k}{m}} \quad \text{and} \quad \omega_2 = \sqrt{\frac{3k}{m}}. \tag{12.16}$$

The eigenvectors (not normalized) corresponding to these eigenvalues are:

(12.17) For $\lambda = 1$: $y = x$ or $\mathbf{r} = (1, 1)$; for $\lambda = 3$: $y = -x$ or $\mathbf{r} = (1, -1)$.

Thus at frequency ω_1 (with $y = x$), the two masses oscillate back and forth together like this $\rightarrow\rightarrow$ and then like this $\leftarrow\leftarrow$. At frequency ω_2 (with $y = -x$), they oscillate in opposite directions like this $\leftarrow\rightarrow$ and then like this $\rightarrow\leftarrow$. These two especially simple ways in which the system can vibrate, each involving just one vibration frequency, are called the characteristic (or normal) modes of vibration; the corresponding frequencies are called the characteristic (or normal) frequencies of the system.

The problem we have just done shows an important method which can be used in many different applications. There are numerous examples of vibration problems in physics—in acoustics: the vibrations of strings of musical instruments, of drumheads, of the air in organ pipes or in a room; in mechanics and its engineering applications: vibrations of mechanical systems all the way from the simple pendulum to complicated structures like bridges and airplanes; in electricity: the vibrations of radio waves, of electric currents and voltages as in a tuned radio; and so on. In such problems, it is often useful to find the characteristic vibration frequencies of the system under consideration and the characteristic modes of vibration. More complicated vibrations can then be discussed as combinations of these simpler normal modes of vibration.

▶ **Example 4.** In Example 3 and Figure 12.1, the two masses were equal and all the spring constants were the same. Changing the spring constants to different values doesn't cause any problems but when the masses are different, there is a possible difficulty which we want to discuss. Consider an array of masses and springs as in Figure 12.1 but with the following masses and spring constants: $2k$, $2m$, $6k$, $3m$, $3k$. We want to find the characteristic frequencies and modes of vibration. Following our work in Example 3, we write the potential energy V, find the forces, write the equations of motion, and substitute $\ddot{x} = -\omega^2 x$, and $\ddot{y} = -\omega^2 y$, in order to find the characteristic frequencies. (*Do the details*: Problem 11.)

$$V = \frac{1}{2}2kx^2 + \frac{1}{2}6k(x-y)^2 + \frac{1}{2}3ky^2 = \frac{1}{2}k(8x^2 - 12xy + 9y^2) \tag{12.18}$$

$$\begin{cases} 2m\ddot{x} = -\partial V/\partial x, \\ 3m\ddot{y} = -\partial V/\partial y, \end{cases} \quad \text{or} \quad \begin{cases} -2m\omega^2 x = \quad -k(8x - 6y), \\ -3m\omega^2 y = -k(-6x + 9y). \end{cases} \tag{12.19}$$

Next divide each equation by its mass and write the equations in matrix form.

$$\omega^2 \begin{pmatrix} x \\ y \end{pmatrix} = \frac{k}{m} \begin{pmatrix} 4 & -3 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{12.20}$$

With $\lambda = m\omega^2/k$, the eigenvalues of the square matrix are $\lambda = 1$ and $\lambda = 6$. Thus the characteristic frequencies of vibration are

$$\omega_1 = \sqrt{\frac{k}{m}} \quad \text{and} \quad \omega_2 = \sqrt{\frac{6k}{m}}. \tag{12.21}$$

FÊUE WHEP

The corresponding eigenvectors are:

(12.22) For $\lambda = 1$: $y = x$ or $\mathbf{r} = (1, 1)$; for $\lambda = 6$: $3y = -2x$ or $\mathbf{r} = (3, -2)$.

Thus at frequency ω_1 the two masses oscillate back and forth together with equal amplitudes like this $\leftarrow\leftarrow$ and then like this $\rightarrow\rightarrow$. At frequency ω_2 the two masses oscillate in opposite directions with amplitudes in the ratio 3 to 2 like this $\leftarrow\rightarrow$ and then like this $\rightarrow\leftarrow$.

Now we seem to have solved the problem; where is the difficulty? Note that the square matrix in (12.20) is not symmetric [and compare (12.14) where the square matrix was symmetric]. In Section 11 we discussed the fact that (for real matrices) only symmetric matrices have orthogonal eigenvectors and can be diagonalized by an orthogonal transformation. Here note that the eigenvectors in Example 3 were orthogonal [dot product of $(1, 1)$ and $(1, -1)$ is zero] but the eigenvectors for (12.20) are not orthogonal [dot product of $(1, 1)$ and $(3, -2)$ is not zero]. If we want orthogonal eigenvectors, we can make the change of variables (also see Example 6)

$$X = x\sqrt{2}, \quad Y = y\sqrt{3}, \tag{12.23}$$

where the constants are the square roots of the numerical factors in the masses $2m$ and $3m$. (Note that geometrically this just amounts to different changes in scale along the two axes, not to a rotation.) Then (12.20) becomes

$$\omega^2 \begin{pmatrix} X \\ Y \end{pmatrix} = \frac{k}{m} \begin{pmatrix} 4 & -\sqrt{6} \\ -\sqrt{6} & 3 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}. \tag{12.24}$$

By inspection we see that the characteristic equation for the square matrix in (12.24) is the same as the characteristic equation for (12.20) so the eigenvalues and the characteristic frequencies are the same as before (as they must be by physical reasoning). However the (12.24) matrix is symmetric and so we know that its eigenvectors are orthogonal. By direct substitution of (12.23) into (12.22), [or by solving for the eigenvectors in the (12.24) matrix] we find the eigenvectors in the X, Y coordinates:

(12.25) For $\lambda = 1$: $\mathbf{R} = (X, Y) = (\sqrt{2}, -\sqrt{3})$; for $\lambda = 6$: $\mathbf{R} = (3\sqrt{2}, 2\sqrt{3}\,)$.

As expected, these eigenvectors are orthogonal.

Example 5. Let's consider a model of a linear triatomic molecule in which we approximate the forces between the atoms by forces due to springs (Figure 12.2).

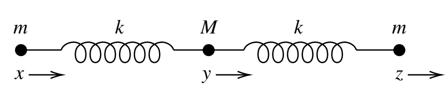


Figure 12.2

As in Example 3, let x, y, z be the coordinates of the three masses relative to their equilibrium positions. We want to find the characteristic vibration frequencies of

10410

the molecule. Following our work in Examples 3 and 4, we find (Problem 12)

$$V = \tfrac{1}{2}k(x-y)^2 + \tfrac{1}{2}k(y-z)^2 = \tfrac{1}{2}k(x^2+2y^2+z^2-2xy-2yz), \tag{12.26}$$

$$\begin{cases} m\ddot{x} = -\partial V/\partial x = -k(x-y), \\ M\ddot{y} = -\partial V/\partial y = -k(2y-x-z), \\ m\ddot{z} = -\partial V/\partial z = -k(z-y), \end{cases}$$

$$\text{or} \tag{12.27}$$

$$\begin{cases} -m\omega^2 x = -k(x-y), \\ -M\omega^2 y = -k(2y-x-z), \\ -m\omega^2 z = -k(z-y). \end{cases}$$

We are going to consider several different ways of solving this problem in order to learn some useful techniques. First of all, if we add the three equations we get

$$m\ddot{x} + M\ddot{y} + m\ddot{z} = 0. \tag{12.28}$$

Physically (12.28) says that the center of mass is at rest or moving at constant speed (that is, has zero acceleration). Since we are just interested in vibrational motion, let's assume that the center of mass is at rest at the origin. Then we have $mx + My + mz = 0$. Solving this equation for y gives

$$y = -\frac{m}{M}(x+z). \tag{12.29}$$

Substitute (12.29) into the second set of equations in (12.27) to get the x and z equations

$$\begin{aligned} -m\omega^2 x &= -k\left(1+\frac{m}{M}\right)x - k\frac{m}{M}z, \\ -m\omega^2 z &= -k\frac{m}{M}x - k\left(1+\frac{m}{M}\right)z. \end{aligned} \tag{12.30}$$

In matrix form equations (12.30) become [compare (12.14)]

$$\lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1+\frac{m}{M} & \frac{m}{M} \\ \frac{m}{M} & 1+\frac{m}{M} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with} \quad \lambda = \frac{m\omega^2}{k}. \tag{12.31}$$

We solve this eigenvalue problem to find

$$\omega_1 = \sqrt{\frac{k}{m}}, \qquad \omega_2 = \sqrt{\frac{k}{m}\left(1+\frac{2m}{M}\right)}. \tag{12.32}$$

For ω_1 we find $z = -x$, and consequently by (12.29), $y = 0$. For ω_2, we find $z = x$ and so $y = -\frac{2m}{M}x$. Thus at frequency ω_1, the central mass M is at rest and the two masses m vibrate in opposite directions like this $\leftarrow m \;\; M \;\; m\rightarrow$ and then like this $m\rightarrow \;\; M \;\; \leftarrow m$. At the higher frequency ω_2, the central mass M moves in one direction while the two masses m move in the opposite direction, first like this $m\rightarrow \;\leftarrow M \;\; m\rightarrow$ and then like this $\leftarrow m \;\; M\rightarrow \;\leftarrow m$.

Now suppose that we had not thought about eliminating the translational motion and had set this problem up as a 3 variable problem. Let's go back to the second set

10411

of equations in (12.27), and divide the x and z equations by m and the y equation by M. Then in matrix form these equations can be written as

$$\omega^2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{k}{m} \begin{pmatrix} 1 & -1 & 0 \\ \frac{-m}{M} & \frac{2m}{M} & \frac{-m}{M} \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{12.33}$$

With $\lambda = m\omega^2/k$, the eigenvalues of the square matrix are $\lambda = 0, 1, 1 + \frac{2m}{M}$, and the corresponding eigenvectors are (check these)

$$\begin{aligned} &\text{For } \lambda = 0,\ \mathbf{r} = (1,1,1); \\ &\text{for } \lambda = 1,\ \mathbf{r} = (1,0,-1); \\ &\text{for } \lambda = 1 + \frac{2m}{M},\ \mathbf{r} = (1, -\frac{2m}{M}, 1). \end{aligned} \tag{12.34}$$

We recognize the $\lambda = 0$ solution as corresponding to translation both because $\omega = 0$ (so there is no vibration), and because $\mathbf{r} = (1,1,1)$ says that any motion is the same for all three masses. The other two modes of vibration are the same ones we had above. We note that the square matrix in (12.33) is not symmetric and so, as expected, the eigenvectors in (12.34) are not an orthogonal set. However, the last two (which correspond to vibrations) are orthogonal so if we are just interested in modes of vibration we can ignore the translation eigenvector. If we want to consider all motion of the molecule along its axis (both translation and vibration), and want an orthogonal set of eigenvectors, we can make the change of variables discussed in Example 4, namely

$$X = x, \quad Y = y\sqrt{\frac{M}{m}}, \qquad Z = z. \tag{12.35}$$

Then the eigenvectors become

$$(1, \sqrt{M/m}, 1), \quad (1, 0, -1), \quad (1, -2\sqrt{m/M}, 1) \tag{12.36}$$

which are an orthogonal set. The first eigenvector (corresponding to translation) may seem confusing, looking as if the central mass M doesn't move with the others (as it must for pure translation). But remember from Example 4 that changes of variable like (12.23) and (12.35) correspond to changes of scale, so in the XYZ system we are not using the same measuring stick to find the position of the central mass as for the other two masses. Their physical displacements are actually all the same.

► **Example 6.** Let's consider Example 4 again in order to illustrate a very compact form for the eigenvalue equation. Satisfy yourself (Problem 13) that we can write the potential energy V in (12.18) as

$$V = \frac{1}{2}k\mathrm{r}^{\mathrm{T}}\mathrm{V}\mathrm{r} \quad \text{where} \quad V = \begin{pmatrix} 8 & -6 \\ -6 & 9 \end{pmatrix}, \quad \mathrm{r} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathrm{r}^{\mathrm{T}} = \begin{pmatrix} x & y \end{pmatrix}. \tag{12.37}$$

Similarly the kinetic energy $T = \frac{1}{2}(2m\dot{x}^2 + 3m\dot{y}^2)$ can be written as

$$T = \frac{1}{2}m\dot{\mathrm{r}}^{\mathrm{T}}\mathrm{T}\dot{\mathrm{r}} \quad \text{where} \quad \mathrm{T} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad \dot{\mathrm{r}} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}, \quad \dot{\mathrm{r}}^{\mathrm{T}} = \begin{pmatrix} \dot{x} & \dot{y} \end{pmatrix}. \tag{12.38}$$

10412

(Notice that the T matrix is diagonal and is a unit matrix when the masses are equal; otherwise T has the mass factors along the main diagonal and zeros elsewhere.) Now using the matrices T and V, we can write the equations of motion (12.19) as

$$m\omega^2\begin{pmatrix}2&0\\0&3\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=k\begin{pmatrix}8&-6\\-6&9\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\quad\text{or}$$

$$\lambda\mathrm{Tr}=\mathrm{Vr}\quad\text{where}\quad\lambda=\frac{m\omega^2}{k}. \tag{12.39}$$

We can think of (12.39) as the basic eigenvalue equation. If T is a unit matrix, then we just have $\lambda\mathrm{r}=\mathrm{Vr}$ as in (12.14). If not, then we can multiply (12.39) by T^{-1} to get

$$\lambda\mathrm{r}=\mathrm{T}^{-1}\mathrm{Vr}=\begin{pmatrix}1/2&0\\0&1/3\end{pmatrix}\begin{pmatrix}8&-6\\-6&9\end{pmatrix}\mathrm{r}=\begin{pmatrix}4&-3\\-2&3\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} \tag{12.40}$$

as in (12.20). However, we see that this matrix is not symmetric and so the eigenvectors will not be orthogonal. If we want the eigenvectors to be orthogonal as in (12.23), we choose new variables so that the T matrix is the unit matrix, that is variables X and Y so that

$$T=\tfrac{1}{2}(2m\dot{x}^2+3m\dot{y}^2)=\tfrac{1}{2}m(\dot{X}^2+\dot{Y}^2). \tag{12.41}$$

But this means that we want $X^2=2x^2$ and $Y^2=3y^2$ as in (12.23), or in matrix form,

$$\mathrm{R}=\begin{pmatrix}X\\Y\end{pmatrix}=\begin{pmatrix}x\sqrt{2}\\y\sqrt{3}\end{pmatrix}=\begin{pmatrix}\sqrt{2}&0\\0&\sqrt{3}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\mathrm{T}^{1/2}\mathrm{r}\quad\text{or}$$

$$\mathrm{r}=\mathrm{T}^{-1/2}\mathrm{R}=\begin{pmatrix}1/\sqrt{2}&0\\0&1/\sqrt{3}\end{pmatrix}\begin{pmatrix}X\\Y\end{pmatrix}. \tag{12.42}$$

Substituting (12.42) into (12.39), we get $\lambda\mathrm{TT}^{-1/2}\mathrm{R}=\mathrm{VT}^{-1/2}\mathrm{R}$. Then multiplying on the left by $\mathrm{T}^{-1/2}$ and noting that $\mathrm{T}^{-1/2}\mathrm{TT}^{-1/2}=\mathrm{I}$, we have

$$\lambda\mathrm{R}=\mathrm{T}^{-1/2}\mathrm{VT}^{-1/2}\mathrm{R} \tag{12.43}$$

as the eigenvalue equation in terms of the new variables X and Y. Substituting the numerical $\mathrm{T}^{-1/2}$ from (12.42) into (12.43) gives the result we had in (12.24).

We have simply demonstrated that (12.39) and (12.43) give compact forms of the eigenvalue equations for Example 4. However, it is straightforward to show that these equations are just a compact summary of the equations of motion for any similar vibrations problem, in any number of variables, just by writing the potential and kinetic energy matrices and comparing the equations of motion in matrix form.

▶ **Example 7.** Find the characteristic frequencies and the characteristic modes of vibration for the system of masses and springs shown in Figure 12.3, where the motion is along a vertical line.

Let's use the simplified method of Example 6 for this problem. We first write the expressions for the kinetic energy and the potential energy as in previous examples.

FEUE WHEP

(Note carefully that we measure x and y from the equilibrium positions of the masses when they are hanging at rest; then the gravitational forces are already balanced and gravitational potential energy does not come into the expression for V.)

$$(12.44)\qquad \begin{aligned} T &= \tfrac{1}{2}m(4\dot{x}^2 + \dot{y}^2), \\ V &= \tfrac{1}{2}k[3x^2 + (x-y)^2] = \tfrac{1}{2}k(4x^2 - 2xy + y^2). \end{aligned}$$

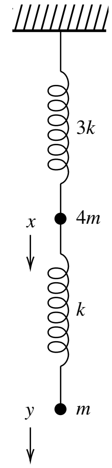


Figure 12.3

The corresponding matrices are [see equations (12.37) and (12.38)]:

$$(12.45)\qquad \mathrm{T} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathrm{V} = \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}.$$

As in equation (12.40), we find $\mathrm{T}^{-1}\mathrm{V}$ and its eigenvalues and eigenvectors.

$$\mathrm{T}^{-1}\mathrm{V} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1/4 \\ -1 & 1 \end{pmatrix}, \qquad \lambda = \frac{m\omega^2}{k} = \frac{1}{2}, \frac{3}{2}.$$

$$(12.46)\qquad \text{For } \omega = \sqrt{\frac{k}{2m}},\ \mathbf{r} = (1,2); \qquad \text{for } \omega = \sqrt{\frac{3k}{2m}},\ \mathbf{r} = (1,-2).$$

As expected (since $\mathrm{T}^{-1}\mathrm{V}$ is not symmetric), the eigenvectors are not orthogonal. If we want orthogonal eigenvectors, we make the change of variables $X = 2x$, $Y = y$, to find the eigenvectors $\mathbf{R} = (1,1)$ and $\mathbf{R} = (1,-1)$ which are orthogonal. Alternatively, we can find the matrix $\mathrm{T}^{-1/2}\mathrm{V}\mathrm{T}^{-1/2}$

$$(12.47)\qquad \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix},$$

and find its eigenvalues and eigenvectors.

PROBLEMS, SECTION 12

1. Verify that (12.2) multiplied out is (12.1).

Find the equations of the following conics and quadric surfaces relative to principal axes.

2. $2x^2 + 4xy - y^2 = 24$

3. $8x^2 + 8xy + 2y^2 = 35$

4. $3x^2 + 8xy - 3y^2 = 8$

5. $5x^2 + 3y^2 + 2z^2 + 4xz = 14$

6. $x^2 + y^2 + z^2 + 4xy + 2xz - 2yz = 12$

7. $x^2 + 3y^2 + 3z^2 + 4xy + 4xz = 60$

8. Carry through the details of Example 2 to find the unit eigenvectors. Show that the resulting rotation matrix C is orthogonal. *Hint:* Find $\mathrm{CC^T}$.

9. For Problems 2 to 7, find the rotation matrix C which relates the principal axes and the original axes. See Example 2.

10. Verify equations (12.13) and (12.14). Solve (12.15) to find the eigenvalues and verify (12.16). Find the corresponding eigenvectors as stated in (12.17).

FEUE WHEP
10414

11. Verify the details of Example 4, equations (12.18) to (12.25).

12. Verify the details of Example 5, equations (12.26) to (12.36).

13. Verify the details of Example 6, equations (12.37) to (12.43).

Find the characteristic frequencies and the characteristic modes of vibration for systems of masses and springs as in Figure 12.1 and Examples 3, 4, and 6 for the following arrays.

14. $k, m, 2k, m, k$

15. $5k, m, 2k, m, 2k$

16. $4k, m, 2k, m, k$

17. $3k, 3m, 2k, 4m, 2k$

18. $2k, m, k, 5m, 10k$

19. $4k, 2m, k, m, k$

20. Carry through the details of Example 7.

Find the characteristic frequencies and the characteristic modes of vibration as in Example 7 for the following arrays of masses and springs, reading from top to bottom in a diagram like Figure 12.3.

21. $3k, m, 2k, m$

22. $4k, 3m, k, m$

23. $2k, 4m, k, 2m$

▶ 13. A BRIEF INTRODUCTION TO GROUPS

We will not go very far into group theory—there are whole books on the subject as well as on its applications in physics. But since so many of the ideas we are discussing in this chapter are involved, it is interesting to have a quick look at groups.

▶ **Example 1.** Think about the four numbers ± 1, $\pm i$. Notice that no matter what products and powers of them we compute, we never get any numbers besides these four. This property of a set of elements with a law of combination is called *closure*. Now think about these numbers written in polar form: $e^{i\pi/2}$, $e^{i\pi}$, $e^{3i\pi/2}$, $e^{2i\pi} = 1$, or the corresponding rotations of a vector (in the xy plane with tail at the origin), or the set of rotation matrices corresponding to these successive 90° rotations of a vector (Problem 1). Note also that these numbers are the four fourth roots of 1, so we could write them as A, A^2, A^3, $A^4 = 1$. All these sets are examples of groups, or more precisely, they are all *representations* of the same group known as the *cyclic group of order* 4. We will be particularly interested in groups of matrices, that is, in matrix representations of groups, since this is very important in applications. Now just what is a group?

Definition of a Group A group is a set $\{A, B, C, \cdots\}$ of elements—which may be numbers, matrices, operations (such as the rotations above)—together with a law of combination of two elements (often called the "product" and written as AB—see discussion below) subject to the following four conditions.

1. Closure: The combination of any two elements is an element of the group.
2. Associative law: The law of combination satisfies the associative law: $(AB)C = A(BC)$.
3. Unit element: There is a unit element I with the property that $IA = AI = A$ for every element of the group.

FEUE WHEP

4. Inverses: Every element of the group has an inverse in the group; that is, for any element A there is an element B such that $AB = BA = I$.

We can easily verify that these four conditions are satisfied for the set ± 1, $\pm i$ under multiplication.

1. We have already discussed closure.

2. Multiplication of numbers is associative.

3. The unit element is 1.

4. The numbers i and $-i$ are inverses since their product is 1; -1 is its own inverse, and 1 is its own inverse.

Thus the set ± 1, $\pm i$, under the operation of multiplication, is a group. The *order of a finite group* is the number of elements in the group. When the elements of a group of order n are of the form A, A^2, A^3, $\cdots$, $A^n = 1$, it is called a *cyclic group*. Thus the group ± 1, $\pm i$, under multiplication, is a cyclic group of order 4 as we claimed above.

A *subgroup* is a subset which is itself a group. The whole group, or the unit element, are called *trivial subgroups*; any other subgroup is called a *proper subgroup*. The group ± 1, $\pm i$ has the proper subgroup ± 1.

Product, Multiplication Table In the definition of a group and in the discussion so far, we have used the term "product" and have written AB for the combination of two elements. However, terms like "product" or "multiplication" are used here in a generalized sense to refer to whatever the operation is for combining group elements. In applications, group elements are often matrices and the operation is matrix multiplication. In general mathematical group theory, the operation might be, for example, addition of two elements, and that sounds confusing to say "product" when we mean sum! Look at one of the first examples we discussed, namely the rotation of a vector by angles $\pi/2$, π, $3\pi/2$, 2π or 0. If the group elements are rotation matrices, then we multiply them, but if the group elements are the angles, then we add them. But the physical problem is exactly the same in both cases. So remember that group multiplication refers to the law of combination for the group rather than just to ordinary multiplication in arithmetic.

Multiplication tables for groups are very useful; equations (13.1), (13.2), and (13.4) show some examples. Look at (13.1) for the group ± 1, $\pm i$. The first column and the top row (set off by lines) list the group elements. The sixteen possible products of these elements are in the body of the table. Note that each element of the group appears exactly once in each row and in each column (Problem 3). At the intersection of the row starting with i and the column headed by $-i$, you find the product $(i)(-i) = 1$, and similarly for the other products.

$$(13.1)\qquad \begin{array}{r|rrrr} & 1 & i & -1 & -i \\ \hline 1 & 1 & i & -1 & -i \\ i & i & -1 & -i & 1 \\ -1 & -1 & -i & 1 & i \\ -i & -i & 1 & i & -1 \end{array}$$

10416

In (13.2) below, note that you add the angles as we discussed above. However, it's not quite just adding—it's really the familiar process of adding angles until you get to 2π and then starting over again at zero. In mathematical language this is called adding (mod 2π) and we write $\pi/2 + 3\pi/2 \equiv 0 \pmod{2\pi}$. Hours on an ordinary clock add in a similar way. If it's 10 o'clock and then 4 hours elapse, the clock says it's 2 o'clock. We write $10 + 4 \equiv 2 \pmod{12}$. (See Problems 6 and 7 for more examples.)

(13.2)

	0	$\pi/2$	π	$3\pi/2$
0	0	$\pi/2$	π	$3\pi/2$
$\pi/2$	$\pi/2$	π	$3\pi/2$	0
π	π	$3\pi/2$	0	$\pi/2$
$3\pi/2$	$3\pi/2$	0	$\pi/2$	π

Two groups are called *isomorphic* if their multiplication tables are identical except for the names we attach to the elements [compare (13.1) and (13.2)]. Thus all the 4-element groups we have discussed so far are isomorphic to each other, that is, they are really all the same group. However, there are two different groups of order 4, the cyclic group we have discussed, and another group called the 4's group (see Problem 4).

Symmetry Group of the Equilateral Triangle Consider three identical atoms at the corners of an equilateral triangle in the xy plane, with the center of the triangle at the origin as shown in Figure 13.1. What rotations and reflections of vectors in the xy plane (as in Section 7) will produce an identical array of atoms? By considering Figure 13.1, we see that there are three possible rotations: $0°$, $120°$, $240°$, and three possible reflections, through the three lines F, G, H (lines along the altitudes of the triangle). Think of moving just the triangle (that is, the atoms), leaving the axes and the lines F, G, H fixed in the background. As in Section 7, we can write a 2 by 2 rotation or reflection matrix for each of these six transformations and set up a multiplication table to show that they do form a group of order 6. This group is called the symmetry group of the equilateral triangle. We find (Problem 8)

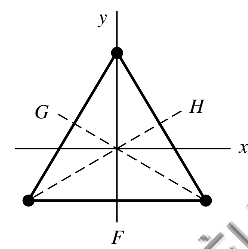


Figure 13.1

(13.3)

Identity, $0°$ rotation $\quad \mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$120°$ rotation $\quad \mathrm{A} = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$

$240°$ rotation $\quad \mathrm{B} = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$

Reflection through line F (y axis) $\quad \mathrm{F} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

Reflection through line G $\quad \mathrm{G} = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$

Reflection through line H $\quad \mathrm{H} = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$

10417

The group multiplication table is:

(13.4)

	I	A	B	F	G	H
I	I	A	B	F	G	H
A	A	B	I	G	H	F
B	B	I	A	H	F	G
F	F	H	G	I	B	A
G	G	F	H	A	I	B
H	H	G	F	B	A	I

Note here that GF = A, but FG = B, not surprising since we know that matrices don't always commute. In group theory, if every two group elements commute, the group is called *Abelian*. Our previous group examples have all been Abelian, but the group in (13.4) is not Abelian.

This is just one example of a symmetry group. Group theory is so important in applications because it offers a systematic way of using the symmetry of a physical problem to simplify the solution. As we have seen, groups can be represented by sets of matrices, and this is widely used in applications.

Conjugate Elements, Class, Character Two group elements A and B are called *conjugate* elements if there is a group element C such that $C^{-1}AC = B$. By letting C be successively one group element after another, we can find all the group elements conjugate to A. This set of conjugate elements is called a *class*. Recall from Section 11 that if A is a matrix describing a transformation (such as a rotation or some sort of mapping of a space onto itself), then $B = C^{-1}AC$ describes the same mapping but relative to a different set of axes (different basis). Thus all the elements of a class really describe the same mapping, just relative to different bases.

Example 2. Find the classes for the group in (13.3) and (13.4). We find the elements conjugate to F as follows [use (13.4) to find inverses and products]:

$$\begin{aligned} I^{-1}FI &= F; \\ A^{-1}FA &= BFA = BH = G; \\ B^{-1}FB &= AFB = AG = H; \\ F^{-1}FF &= F; \\ G^{-1}FG &= GFG = GB = H; \\ H^{-1}FH &= HFH = HA = G. \end{aligned} \tag{13.5}$$

Thus the elements F, G, and H are conjugate to each other and form one class. You can easily show (Problem 12) that elements A and B are another class, and the unit element I is a class by itself. Now notice what we observed above. The elements F, G, and H all just interchange two atoms, that is, all of them do the same thing, just seen from a different viewpoint. The elements A and B rotate the atoms, A by 120° and B by 240° which is the same as 120° looked at upside down. And finally the unit element I leaves things unchanged so it is a class by itself. Notice that a class is not a group (except for the class consisting of I) since a group must contain the unit element. So a class is a subset of a group, but not a subgroup.

10418

Recall from (9.13) and Problem 11.10 that the trace of a matrix (sum of diagonal elements) is not changed by a similarity transformation. Thus all the matrices of a class have the same trace. Observe that this is true for the group (13.3): Matrix I has trace = 2, A and B have trace $= -\frac{1}{2} - \frac{1}{2} = -1$, and F, G, and H have trace = 0. In this connection, the trace of a matrix is called its *character*, so we see that all matrices of a class have the same character. Also note that we could write the matrices (13.3) in (infinitely) many other ways by rotating the reference axes, that is, by performing similarity transformations. But since similarity transformations do not change the trace, that is, the character, we now have a number attached to each class which is independent of the particular choice of coordinate system (basis). Classes and their associated character are very important in applications of group theory.

One more number is important here, and that is the dimension of a representation. In (13.3), we used 2 by 2 matrices (2 dimensions), but it would be possible to work in 3 dimensions. Then, for example, the A matrix would describe a 120° rotation around the z axis and would be

$$\text{A} = \begin{pmatrix} -1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{13.6}$$

and the other matrices in (13.3) would have a similar form, called *block diagonalized*. But now the traces of all the matrices are increased by 1. To avoid having any ambiguity about character, we use what are called "irreducible representations" in finding character; let's discuss this.

Irreducible Representations A 2-dimensional representation is called *reducible* if all the group matrices can be diagonalized by the same unitary similarity transformation (that is, the same change of basis). For example, the matrices in Problem 1 and the matrices in Problem 4 both give 2-dimensional reducible representations of their groups (see Problems 13, 15, and 16). On the other hand, the matrices in (13.3) cannot be simultaneously diagonalized (see Problem 13), so (13.3) is called a 2-dimensional *irreducible representation* of the equilateral triangle symmetry group. If a group of 3 by 3 matrices can all be either diagonalized or put in the form of (13.6) (block diagonalized) by the same unitary similarity transformation, then the representation is called reducible; if not, it is a 3-dimensional irreducible representation. For still larger matrices, imagine the matrices block diagonalized with blocks along the main diagonal which are the matrices of irreducible representations.

Thus we see that any representation is made up of irreducible representations. For each irreducible representation, we find the character of each class. Such lists are known as character tables, but their construction is beyond our scope.

Infinite Groups Here we survey some examples of infinite groups as well as some sets which are not groups.

(13.7)

(a) The set of all integers, positive, negative, and zero, under ordinary addition, is a group. *Proof*: The sum of two integers is an integer. Ordinary addition obeys the associative law. The unit element is 0. The inverse of the integer N is −N since N + (−N) = 0.

FÊUE WHEP

(b) The same set under ordinary multiplication is not a group because 0 has no inverse. But even if we omit 0, the inverses of the other integers are fractions which are not in the set.

(c) Under ordinary multiplication, the set of all rational numbers except zero, is a group. *Proof*: The product of two rational numbers is a rational number. Ordinary multiplication is associative. The unit element is 1, and the inverse of a rational number is just its reciprocal.

Similarly, you can show that the following sets are groups under ordinary multiplication (Problem 17): All real numbers except zero, all complex numbers except zero, all complex numbers $re^{i\theta}$ with $r = 1$.

(d) Ordinary subtraction or division cannot be group operations because they don't satisfy the associative law; for example, $x-(y-z) \neq (x-y)-z$. (Problem 18.)

(e) The set of all orthogonal 2 by 2 matrices under matrix multiplication is a group called O(2). If the matrices are required to be rotation matrices, that is, have determinant +1, the set is a group called SO(2) (the S stands for special). Similarly, the following sets of matrices are groups under matrix multiplication: The set of all orthogonal 3 by 3 matrices, called O(3); its subgroup SO(3) with determinant = 1; or the corresponding sets of orthogonal matrices of any dimension n, called O(n) and SO(n). (Problem 19.)

(f) The set of all unitary n by n matrices, $n = 1, 2, 3, \cdots$, called U(n), is a group under matrix multiplication, and its subgroup SU(n) of unitary matrices with determinant = 1 is also a group. *Proof*: We have repeatedly noted that matrix multiplication is associative and that the unit matrix is the unit element of a group of matrices. So we just need to check closure and inverses. The product of two unitary matrices is unitary (see Section 9). If two matrices have determinant = 1, their product has determinant = 1 [see equation (6.6)]. The inverse of a unitary matrix is unitary (see Problem 9.25).

PROBLEMS, SECTION 13

1. Write the four rotation matrices for rotations of vectors in the xy plane through angles 90°, 180°, 270°, 360° (or 0°) [see equation (7.12)]. Verify that these 4 matrices under matrix multiplication satisfy the four group requirements and are a matrix representation of the cyclic group of order 4. Write their multiplication table and compare with Equations (13.1) and (13.2).

2. Following the text discussion of the cyclic group of order 4, and Problem 1, discuss

(a) the cyclic group of order 3 (see Chapter 2, Problem 10.32);

(b) the cyclic group of order 6.

3. Show that, in a group multiplication table, each element appears exactly once in each row and in each column. *Hint:* Suppose that an element appears twice, and show that this leads to a contradiction, namely that two elements assumed different are the same element.

4. Show that the matrices

$$\mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad \mathrm{C} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

10420

under matrix multiplication, form a group. Write the group multiplication table to see that this group (called the 4's group) is not isomorphic to the cyclic group of order 4 in Problem 1. Show that the 4's group is Abelian but not cyclic.

5. Consider the group of order 4 with unit element I and other elements A, B, C, where $AB = BA = C$, and $A^2 = B^2 = I$. Write the group multiplication table and verify that it is a group. There are two groups of order 4 (discussed in Problems 1 and 4). To which is this one isomorphic? *Hint:* Compare the multiplication tables.

6. Consider the integers 0, 1, 2, 3 under addition (mod 4). Write the group "multiplication" table and show that you have a group of order 4. Is this group isomorphic to the cyclic group of order 4 or to the 4's group?

7. Consider the set of numbers 1, 3, 5, 7 with multiplication (mod 8) as the law of combination. Write the multiplication table to show that this is a group. [To multiply two numbers (mod 8), you multiply them and then take the remainder after dividing by 8. For example, $5 \times 7 = 35 \equiv 3 (\text{mod } 8)$.] Is this group isomorphic to the cyclic group of order 4 or to the 4's group?

8. Verify (13.3) and (13.4). *Hints:* For the rotation and reflection matrices, see Section 7. In checking the multiplication table, be sure you are multiplying the matrices in the right order. Remember that matrices are operators on the vectors in the plane (Section 7), and matrices may not commute. GFA means apply A, then F, then G.

9. Show that any cyclic group is Abelian. *Hint:* Does a matrix commute with itself?

10. As we did for the equilateral triangle, find the symmetry group of the square. *Hints:* Draw the square with its center at the origin and its sides parallel to the x and y axes. Find a set of eight 2 by 2 matrices (4 rotation and 4 reflection) which map the square onto itself, and write the multiplication table to show that you have a group.

11. Do Problem 10 for a rectangle. Note that now only two rotations and 2 reflections leave the rectangle unchanged. So you have a group of order 4. To which is it isomorphic, the cyclic group or the 4's group?

12. Verify (13.5) and then also show that A, B are the elements of a class, and that I is a class by itself. Show that it will always be true in any group that I is a class by itself. *Hint:* What is $\mathrm{C}^{-1}\mathrm{IC}$ for any element C of a group?

13. Using the discussion of simultaneous diagonalization at the end of Section 11, show that the 2-dimensional matrices in Problems 1 and 4 are reducible representations of their groups, and the matrices in (13.5) give an irreducible representation of the equilateral triangle symmetry group. *Hint:* Look at the multiplication tables to see which matrices commute.

14. Use the multiplication table you found in Problem 10 to find the classes in the symmetry group of a square. Show that the 2 by 2 matrices you found are an irreducible representation of the group (see Problem 13), and find the character of each class for that representation. Note that it is possible for the character to be the same for two classes, but it is not possible for the character of two elements of the same class to be different.

15. By Problem 13, you know that the matrices in Problem 4 are a reducible representation of the 4's group, that is they can all be diagonalized by the same unitary similarity transformation (in this case orthogonal since the matrices are symmetric). Demonstrate this directly by finding the matrix C and diagonalizing all 4 matrices.

16. Do Problem 15 for the group of matrices you found in Problem 1. Be careful here—you are working in a complex vector space and your C matrix will be unitary but

FEUE WHEP

not orthogonal (see Sections 10 and 11). *Comment:* Not surprisingly, the numbers 1, i, -1, $-i$ give a 1-dimensional representation—note that a single number can be thought of as a 1-dimensional matrix.

17. Verify that the sets listed in (13.7c) are groups.

18. Show that division cannot be a group operation. *Hint:* See (13.7d).

19. Verify that the sets listed in (13.7e) are groups. *Hint:* See the proofs in (13.7f).

20. Is the set of all orthogonal 3-by-3 matrices with determinant $= -1$ a group? If so, what is the unit element?

21. Is the group SO(2) Abelian? What about SO(3)? *Hint:* See the discussion following equation (6.14).

14. GENERAL VECTOR SPACES

In this section we are going to introduce a generalization of our picture of vector spaces which is of great importance in applications. This will be merely an introduction because the ideas here will be used in many of the following chapters as you will discover. The basic idea will be to set up an outline of the requirements for 3-dimensional vector spaces (as we listed the requirements for a group), and then show that these familiar 3-dimensional vector space requirements are satisfied by sets of things like functions or matrices which we would not ordinarily think of as vectors.

Definition of a Vector Space A vector space is a set of elements $\{\mathbf{U}, \mathbf{V}, \mathbf{W}, \cdots\}$ called vectors, together with two operations: addition of vectors, and multiplication of a vector by a scalar (which for our purposes will be a real or a complex number), and subject to the following requirements:

1. Closure: The sum of any two vectors is a vector in the space.
2. Vector addition is:
   (a) commutative: $\mathbf{U}+\mathbf{V}=\mathbf{V}+\mathbf{U}$,
   (b) associative: $(\mathbf{U}+\mathbf{V})+\mathbf{W}=\mathbf{U}+(\mathbf{V}+\mathbf{W})$.
3. **(a)** There is a zero vector $\mathbf{0}$ such that $\mathbf{0}+\mathbf{V}=\mathbf{V}+\mathbf{0}=\mathbf{V}$ for every element $\mathbf{V}$ in the space.
   (b) Every element $\mathbf{V}$ has an additive inverse $(-\mathbf{V})$ such that $\mathbf{V}+(-\mathbf{V})=\mathbf{0}$.
4. Multiplication of vectors by scalars has the expected properties:
   (a) $k(\mathbf{U}+\mathbf{V})=k\mathbf{U}+k\mathbf{V}$;
   (b) $(k_1+k_2)\mathbf{V}=k_1\mathbf{V}+k_2\mathbf{V}$;
   (c) $(k_1k_2)\mathbf{V}=k_1(k_2\mathbf{V})$;
   (d) $0\cdot\mathbf{V}=\mathbf{0}$, and $1\cdot\mathbf{V}=\mathbf{V}$.

You should go over these and satisfy yourself that they are all true for ordinary two and three dimensional vector spaces. Now let's look at some examples of things we don't usually think of as vectors which, nevertheless, satisfy the above requirements.

10422

▶ **Example 1.** Consider the set of polynomials of the third degree or less, namely functions of the form $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3$. Is this a vector space? If so, find a basis. What is the dimension of the space?

We go over the requirements listed above:

1. The sum of two polynomials of degree ≤ 3 is a polynomial of degree ≤ 3 and so is a member of the set.

2. Addition of algebraic expressions is commutative and associative.

3. The "zero vector" is the polynomial with all coefficients a_i equal to 0, and adding it to any other polynomial just gives that other polynomial. The additive inverse of a function $f(x)$ is just $-f(x)$, and $-f(x) + f(x) = 0$ as required for a vector space.

4. All the listed familiar rules are just what we do every time we work with algebraic expressions.

So we have a vector space! Now let's try to find a basis for it. Consider the set of functions: $\{1, x, x^2, x^3\}$. They span the space since any polynomial of degree ≤ 3 is a linear combination of them. You can easily show (Problem 1) by computing the Wronskian [equation (8.5)] that they are linearly independent. Therefore they are a basis, and since there are 4 basis vectors, the dimension of the space is 4.

▶ **Example 2.** Consider the set of linear combinations of the functions

$$\{e^{ix},\ e^{-ix},\ \sin x,\ \cos x,\ x\sin x\}.$$

It is straightforward to verify that all our requirements above are met (Problem 1). To find a basis, we must find a linearly independent set of functions which spans the space. We note that the given functions are not linearly independent since e^{ix} and e^{-ix} are linear combinations of $\sin x$ and $\cos x$ (Chapter 2, Section 4). However, the set $\{\sin x,\ \cos x,\ x\sin x\}$ is a linearly independent set and it spans the space. So this is a possible basis and the dimension of the space is 3. Another possible basis would be $\{e^{ix},\ e^{-ix},\ x\sin x\}$. You will meet sets of functions like these as solutions of differential equations (*see* Chapter 8, Problems 5.13 to 5.18).

▶ **Example 3.** Modify Example 1 to consider the set of polynomials of degree ≤ 3 with $f(1) = 1$. Is this a vector space? Suppose we add two of the polynomials; then the value of the sum at $x = 1$ is 2, so it is not an element of the set. Thus requirement 1 is not satisfied so this is not a vector space. Note that a subset of the vectors of a vector space is not necessarily a subspace. On the other hand, if we consider polynomials of degree ≤ 3 with $f(1) = 0$, then the sum of two of them is zero at $x = 1$; this is a vector space. You can easily verify (Problem 1) that it is a subspace of dimension 3 and a possible basis is $\{x-1,\ x^2-1,\ x^3-1\}$.

▶ **Example 4.** Consider the set of all polynomials of any degree $\leq N$. The sum of two polynomials of degree $\leq N$ is another such polynomial, and you can easily verify (Problem 1) that the rest of the requirements are met, so this is a vector space. A simple choice of basis is the set of powers of x from $x^0 = 1$ to x^N. Thus we see that the dimension of this space is $N + 1$.

FEUE WHEP

Example 5. Consider the set of all 2 by 3 matrices with matrix addition as the law of combination, and multiplication by scalars defined as in Section 6. Recall that you add matrices by adding corresponding elements. Thus a sum of two 2 by 3 matrices is another 2 by 3 matrix. For matrix addition and multiplication by scalars, it is straightforward to show that the other requirements listed above are satisfied (Problem 1). As a basis, we could use the six matrices:

$$\begin{pmatrix}1 & 0 & 0\\0 & 0 & 0\end{pmatrix}, \quad \begin{pmatrix}0 & 1 & 0\\0 & 0 & 0\end{pmatrix}, \quad \begin{pmatrix}0 & 0 & 1\\0 & 0 & 0\end{pmatrix},$$
$$\begin{pmatrix}0 & 0 & 0\\1 & 0 & 0\end{pmatrix}, \quad \begin{pmatrix}0 & 0 & 0\\0 & 1 & 0\end{pmatrix}, \quad \begin{pmatrix}0 & 0 & 0\\0 & 0 & 1\end{pmatrix}.$$

Satisfy yourself that these are linearly independent and that they span the space (that is, that you could write any 2 by 3 matrix as a linear combination of these six). Since there are 6 basis vectors, the dimension of this space is 6.

Inner Product, Norm, Orthogonality The definitions of these terms need to be generalized when our "vectors" are functions, that is, we want to generalize equations (10.1) to (10.3). A natural generalization of a sum is an integral, so we might reasonably replace $\sum A_i B_i$ by $\int A(x)B(x)\,dx$, and $\sum A_i^2$ by $\int [A(x)]^2\,dx$. However, in applications we frequently want to consider complex functions of the real variable x (for example, e^{ix} as in Example 2). Thus, given functions $A(x)$ and $B(x)$ on $a \le x \le b$, we define

$$\text{(14.1)} \qquad [\text{Inner Product of } A(x) \text{ and } B(x)] = \int_a^b A^*(x)B(x)\,dx,$$

$$\text{(14.2)} \qquad [\text{Norm of } A(x)] = ||A(x)|| = \sqrt{\int_a^b A^*(x)A(x)\,dx},$$

$$\text{(14.3)} \qquad A(x) \text{ and } B(x) \text{ are orthogonal on } (a,b) \quad \text{if} \quad \int_a^b A^*(x)B(x)\,dx = 0.$$

Let's now generalize our definition (14.1) of inner product still further. Let A, B, C, $\cdots$ be elements of a vector space, and let a, b, c, $\cdots$ be scalars. We will use the bracket $\langle A|B\rangle$ to mean the inner product of A and B. This vector space is called an *inner product space* if an inner product is defined subject to the conditions:

(14.4a) $\qquad \langle A|B\rangle^* = \langle B|A\rangle;$

(14.4b) $\qquad \langle A|A\rangle \ge 0, \quad \langle A|A\rangle = 0$ if and only if $A = 0;$

(14.4c) $\qquad \langle C|aA + bB\rangle = a\langle C|A\rangle + b\langle C|B\rangle.$

(See Problem 11.) It follows from (14.4) that (Problem 12)

(14.5a) $\qquad \langle aA + bB|C\rangle = a^*\langle A|C\rangle + b^*\langle B|C\rangle,$ and

(14.5b) $\qquad \langle aA|bB\rangle = a^* b\langle A|B\rangle.$

You will find various other notations for the inner product, such as (A, B) or $[A, B]$ or $\langle A, B\rangle$. The notation $\langle A|B\rangle$ is used in quantum mechanics. Most mathematics books put the complex conjugate on the second factor in (14.1) and make the corresponding changes in (14.4) and (14.5). Most physics and mathematical methods

FÊUE WHEP

books handle the complex conjugate as we have. If you are confused by this notation and equations (14.4) and (14.5), keep going back to (14.1) where $\langle A|B\rangle = \int A^* B$ until you get used to the bracket notation. Also study carefully our use of the bracket notation in the next section and do Problems 11 to 14.

Schwarz's Inequality In Section 10 we proved the Schwarz inequality for n-dimensional Euclidean space. For an inner product space satisfying (14.4), it becomes [compare (10.9)]

$$|\langle A|B\rangle|^2 \le \langle A|A\rangle\langle B|B\rangle. \tag{14.6}$$

To prove this, we first note that it is true if $B = 0$. For $B \neq 0$, let $C = A - \mu B$, where $\mu = \langle B|A\rangle/\langle B|B\rangle$, and find $\langle C|C\rangle$ which is ≥ 0 by (14.4b). Using (14.4) and (14.5), we write

$$\langle A - \mu B|A - \mu B\rangle = \langle A|A\rangle - \mu^*\langle B|A\rangle - \mu\langle A|B\rangle + \mu^*\mu\langle B|B\rangle \ge 0. \tag{14.7}$$

Now substitute the values of μ and μ^* to get (see Problem 13)

$$\begin{aligned} &\langle A|A\rangle - \frac{\langle A|B\rangle}{\langle B|B\rangle}\langle B|A\rangle - \frac{\langle B|A\rangle}{\langle B|B\rangle}\langle A|B\rangle + \frac{\langle A|B\rangle}{\langle B|B\rangle}\frac{\langle B|A\rangle}{\langle B|B\rangle}\langle B|B\rangle \\ &\qquad = \langle A|A\rangle - \frac{\langle A|B\rangle\langle A|B\rangle^*}{\langle B|B\rangle} = \langle A|A\rangle - \frac{|\langle A|B\rangle|^2}{\langle B|B\rangle} \ge 0 \end{aligned} \tag{14.8}$$

which gives (14.6).

For a function space as in (14.1) to (14.3), Schwarz's inequality becomes (see Problem 14):

$$\left|\int_a^b A^*(x)B(x)\,dx\right|^2 \le \left(\int_a^b A^*(x)A(x)\,dx\right)\left(\int_a^b B^*(x)B(x)\,dx\right). \tag{14.9}$$

Orthonormal Basis; Gram-Schmidt Method Two functions are called *orthogonal* if they satisfy (14.3); a function is *normalized* if its norm in (14.2) is 1. By a combination of the two words, we call a set of functions *orthonormal* if they are all mutually orthogonal and they all have norm 1. It is often convenient to write the functions of a vector space in terms of an orthonormal basis (compare writing ordinary vectors in three dimensions in terms of **i**, **j**, **k**). Let's see how the Gram-Schmidt method applies to a vector space of functions with inner product, norm, and orthogonality defined by (14.1) to (14.3). (Compare Section 10, Example 4 and the paragraph before it.)

▶ **Example 6.** In Example 1, we found that the set of all polynomials of degree ≤ 3 is a vector space of dimension 4 with basis $1, x, x^2, x^3$. Let's consider these polynomials on the interval $-1 \le x \le 1$ and construct an orthonormal basis. To keep track of what we're doing, let $f_0, f_1, f_2, f_3 = 1, x, x^2, x^3$; let p_0, p_1, p_2, p_3 be a corresponding orthogonal basis (which we find by the Gram-Schmidt method); and let e_0, e_1, e_2, e_3, be the orthonormal basis (which we get by normalizing the functions p_i). Recall the Gram-Schmidt routine (see Section 10, Example 4): Normalize the first function

FÊUE WHEP

to get e_0. Then for the rest of the functions, subtract from f_i each preceding e_j multiplied by the inner product of e_j and f_i, that is, find

$$p_i = f_i - \sum_{j<i} e_j \langle e_j | f_i \rangle = f_i - \sum_{j<i} e_j \int_{-1}^{1} e_j f_i \, dx. \tag{14.10}$$

Finally, normalize p_i to get e_i.

We can save effort by noting in advance that many of the inner products we need are going to be zero. You can easily show (Problem 15) that the integral of an odd power of x from $x = -1$ to 1 is zero, and consequently any even power of x is orthogonal to any odd power. Observe that the f_i are alternately even and odd powers of x. Then you can show that the corresponding p_i and e_i will also involve just even or just odd powers of x. The Gram-Schmidt method gives the following results (Problem 16).

$$f_0 = 1 = p_0, \quad ||p_0||^2 = \int_{-1}^{1} 1^2 \, dx = 2, \quad e_0 = \frac{1}{\sqrt{2}}.$$

$f_1 = x; \quad p_1 = x$ because x is orthogonal to e_0.

$$||p_1||^2 = \int_{-1}^{1} x^2 \, dx = \frac{2}{3}, \quad e_1 = x\sqrt{\frac{3}{2}}.$$

$f_2 = x^2$. Since x^2 is orthogonal to e_1 but not to e_0,

$$p_2 = x^2 - \frac{1}{\sqrt{2}} \int_{-1}^{1} \frac{1}{\sqrt{2}} x^2 \, dx = x^2 - \frac{1}{3}.$$

$$||p_2||^2 = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45}, \quad e_2 = (3x^2 - 1)\sqrt{\frac{5}{8}}.$$

$f_3 = x^3$. Since x^3 is orthogonal to e_0 and e_2,

$$p_3 = x^3 - x\sqrt{\frac{3}{2}} \int_{-1}^{1} x\sqrt{\frac{3}{2}} x^3 \, dx = x^3 - \frac{3}{5}x,$$

$$||p_3||^2 = \int_{-1}^{1} \left(x^3 - \frac{3}{5}x\right)^2 dx = \frac{8}{175}, \quad e_3 = (5x^3 - 3x)\sqrt{\frac{7}{8}}.$$

This process could be continued for a vector space with basis 1, x, x^2, $\cdots$, x^N (but it is not very efficient). The orthonormal functions e_i are well-known functions called (normalized) *Legendre polynomials*. In Chapters 12 and 13, we will discover these functions as solutions of differential equations and see their applications in physics problems.

Infinite Dimensional Spaces If a vector space does not have a finite basis, it is called an infinite dimensional vector space. It is beyond our scope to go into a detailed mathematical study of such spaces. However, you should know that, by analogy with finite dimensional vector spaces, we still use the term basis functions for sets of functions (like x^n or $\sin nx$) in terms of which we can expand suitably restricted functions in infinite series. So far we have discussed only power series (Chapter 1). In later chapters you will discover many other sets of functions which provide useful bases in applications: sines and cosines in Chapter 7, various special functions in Chapters 12 and 13. When we introduce them, we will discuss questions of convergence of the infinite series, and of completeness of sets of basis functions.

FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP FEUE WHEP 10426

PROBLEMS, SECTION 14

1. Verify the statements indicated in Examples 1 to 5 above.

For each of the following sets, either verify (as in Example 1) that it is a vector space, or show which requirements are not satisfied. If it is a vector space, find a basis and the dimension of the space.

2. Linear combinations of the set of functions $\{e^x, \sinh x, xe^x\}$.

3. Linear combinations of the set of functions $\{x, \cos x, x\cos x, e^x\cos x, (2-3e^x)\cos x, x(1+5\cos x)\}$.

4. Polynomials of degree ≤ 3 with $a_2 = 0$.

5. Polynomials of degree ≤ 5 with $a_1 = a_3$.

6. Polynomials of degree ≤ 6 with $a_3 = 3$.

7. Polynomials of degree ≤ 7 with all the even coefficients equal to each other and all the odd coefficients equal to each other.

8. Polynomials of degree ≤ 7 but with all odd powers missing.

9. Polynomials of degree ≤ 10 but with all even powers having positive coefficients.

10. Polynomials of degree ≤ 13, but with the coefficient of each odd power equal to half the preceding coefficient of an even power.

11. Verify that the definitions in (14.1) and (14.2) satisfy the requirements for an inner product listed in (14.4) and (14.5). *Hint:* Write out all the equations (14.4) and (14.5) in the integral notation of (14.1) and (14.2).

12. Verify that the relations in (14.5) follow from (14.4). *Hints:* For (14.5a), take the complex conjugate of (14.4c). To take the complex conjugate of a bracket, use (14.4a).

13. Verify (14.7) and (14.8) *Hints:* Remember that a norm squared, like $\langle B|B\rangle$, is a real and non-negative.scalar, so its complex conjugate is just itself. But $\langle B|A\rangle$ is a complex scalar and $\langle B|A\rangle = \langle A|B\rangle^*$ by (14.4). Show that $\mu^* = \langle A|B\rangle/\langle B|B\rangle$.

14. Verify that (14.9) is (14.6) with the definition of scalar product as in (14.1).

15. For Example 6, verify the claimed orthogonality on $(-1,1)$ of an even power of x and an odd power of x. *Hint:* For example, consider $\int_{-1}^{1} x^2x^3\,dx$.

16. For Example 6, verify the details of the terms omitted in the functions p_i because of orthogonality. *Hint:* See Problem 15. Also verify the calculations of inner products and norms and the orthonormal set e_i.

15. MISCELLANEOUS PROBLEMS

1. Show that if each element of one row (or column) of a determinant is the sum of two terms, the determinant can be written as a sum of two determinants; for example,

$$\begin{vmatrix} a_{11} & a_{12}+b_{12} & a_{13} \\ a_{21} & a_{22}+b_{22} & a_{23} \\ a_{31} & a_{32}+b_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & b_{12} & a_{13} \\ a_{21} & b_{22} & a_{23} \\ a_{31} & b_{32} & a_{33} \end{vmatrix}.$$

Use this result to verify Fact 4b of Section 3.

FEUE WHEP
10427

2. What is wrong with the following argument? "If we add the first row of a determinant to the second row and the second row to the first row, then the first two rows of the determinant are identical, and the value of the determinant is zero. Therefore all determinants have the value zero."

3. (a) Find the equations of the line through the points $(4, -1, 2)$ and $(3, 1, 4)$.

(b) Find the equation of the plane through the points $(0, 0, 0)$, $(1, 2, 3)$ and $(2, 1, 1)$.

(c) Find the distance from the point $(1, 1, 1)$ to the plane $3x - 2y + 6z = 12$.

(d) Find the distance from the point $(1, 0, 2)$ to the line $\mathbf{r} = 2\mathbf{i}+\mathbf{j}-\mathbf{k}+(\mathbf{i}-2\mathbf{j}+2\mathbf{k})t$.

(e) Find the angle between the plane in (c) and the line in (d).

4. Given the line $\mathbf{r} = 3\mathbf{i} - \mathbf{j} + (2\mathbf{i} + \mathbf{j} - 2\mathbf{k})t$:

(a) Find the equation of the plane containing the line and the point $(2, 1, 0)$.

(b) Find the angle between the line and the (y, z) plane.

(c) Find the perpendicular distance between the line and the x axis.

(d) Find the equation of the plane through the point $(2, 1, 0)$ and perpendicular to the line.

(e) Find the equations of the line of intersection of the plane in (d) and the plane $y = 2z$.

5. (a) Write the equations of a straight line through the points $(2, 7, -1)$ and $(5, 7, 3)$.

(b) Find the equation of the plane determined by the two lines $\mathbf{r} = (\mathbf{i} - 2\mathbf{j} + \mathbf{k})t$ and $\mathbf{r} = (6\mathbf{i} - 3\mathbf{j} + 2\mathbf{k})t$.

(c) Find the angle which the line in (a) makes with the plane in (b).

(d) Find the distance from $(1, 1, 1)$ to the plane in (b).

(e) Find the distance from $(1, 6, -3)$ to the line in (a).

6. Derive the formula

$$D = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}$$

for the distance from (x_0, y_0, z_0) to $ax + by + cz = d$.

7. Given the matrices A, B, C below, find or mark as meaningless the matrices: A^{T}, A^{-1}, AB, $\bar{\mathrm{A}}$, $\mathrm{A}^{\mathrm{T}}\mathrm{B}^{\mathrm{T}}$, $\mathrm{B}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}$, BA^{T}, ABC, $\mathrm{AB}^{\mathrm{T}}\mathrm{C}$, $\mathrm{B}^{\mathrm{T}}\mathrm{AC}$, $\mathrm{A}^{\dagger}$, $\mathrm{B}^{\mathrm{T}}\mathrm{C}$, $\mathrm{B}^{-1}\mathrm{C}$, $\mathrm{C}^{-1}\mathrm{A}$, CB^{T}.

$$\mathrm{A} = \begin{pmatrix} 1 & -1 \\ 0 & i \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 5 \end{pmatrix}, \quad \mathrm{C} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

8. Given

$$\mathrm{A} = \begin{pmatrix} 1 & 0 & 2i \\ i & -3 & 0 \\ 1 & 0 & i \end{pmatrix}, \quad \text{find } \mathrm{A}^{\mathrm{T}}, \bar{\mathrm{A}}, \mathrm{A}^{\dagger}, \mathrm{A}^{-1}.$$

9. The following matrix product is used in discussing a thick lens in air:

$$\mathrm{A} = \begin{pmatrix} 1 & (n-1)/R_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ d/n & 1 \end{pmatrix} \begin{pmatrix} 1 & -(n-1)/R_1 \\ 0 & 1 \end{pmatrix},$$

where d is the thickness of the lens, n is its index of refraction, and R_1 and R_2 are the radii of curvature of the lens surfaces. It can be shown that element A_{12} of A is $-1/f$ where f is the focal length of the lens. Evaluate A and det A (which should equal 1) and find $1/f$. [See Am. J. Phys. **48**, 397–399 (1980).]

FEUE WHEP
10428

10. The following matrix product is used in discussing two thin lenses in air:

$$\mathrm{M} = \begin{pmatrix} 1 & -1/f_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/f_1 \\ 0 & 1 \end{pmatrix},$$

where f_1 and f_2 are the focal lengths of the lenses and d is the distance between them. As in Problem 9, element M_{12} is $-1/f$ where f is the focal length of the combination. Find M, $\det \mathrm{M}$, and $1/f$.

11. There is a one-to-one correspondence between two-dimensional vectors and complex numbers. Show that the real and imaginary parts of the product $z_1 z_2^*$ (the star denotes complex conjugate) are respectively the scalar product and $\pm$ the magnitude of the vector product of the vectors corresponding to z_1 and z_2.

12. The vectors $\mathbf{A} = a\mathbf{i} + b\mathbf{j}$ and $\mathbf{B} = c\mathbf{i} + d\mathbf{j}$ form two sides of a parallelogram. Show that the area of the parallelogram is given by the absolute value of the following determinant. (Also see Chapter 6, Section 3.)

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

13. The plane $2x+3y+6z=6$ intersects the coordinate axes at points P, Q, R, forming a triangle. Find the vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$. Write a vector formula for the area of the triangle PQR, and find the area.

In Problems 14 to 17, multiply matrices to find the resultant transformation. *Caution:* Be sure you are multiplying the matrices in the right order.

14. $\begin{cases} x' = (x + y\sqrt{3})/2 \\ y' = (-x\sqrt{3} + y)/2 \end{cases}$ $\begin{cases} x'' = (-x' + y'\sqrt{3})/2 \\ y'' = -(x'\sqrt{3} + y')/2 \end{cases}$

15. $\begin{cases} x' = 2x + 5y \\ y' = x + 3y \end{cases}$ $\begin{cases} x'' = x' - 2y' \\ y'' = 3x' - 5y' \end{cases}$

16. $\begin{cases} x' = (x + y\sqrt{2} + z)/2 \\ y' = (x\sqrt{2} - z\sqrt{2})/2 \\ z' = (-x + y\sqrt{2} - z)/2 \end{cases}$ $\begin{cases} x'' = (x'\sqrt{2} + z'\sqrt{2})/2 \\ y'' = (-x' - y'\sqrt{2} + z')/2 \\ z'' = (x' - y'\sqrt{2} - z')/2 \end{cases}$

17. $\begin{cases} x' = (2x + y + 2z)/3 \\ y' = (x + 2y - 2z)/3 \\ z' = (2x - 2y - z)/3 \end{cases}$ $\begin{cases} x'' = (2x' + y' + 2z')/3 \\ y'' = (-x' - 2y' + 2z')/3 \\ z'' = (-2x' + 2y' + z')/3 \end{cases}$

Find the eigenvalues and eigenvectors of the matrices in the following problems.

18. $\begin{pmatrix} 1 & 0 \\ 3 & -2 \end{pmatrix}$ **19.** $\begin{pmatrix} 5 & 1 \\ 4 & 2 \end{pmatrix}$ **20.** $\begin{pmatrix} 5 & -4 \\ -4 & 5 \end{pmatrix}$ **21.** $\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$

22. $\begin{pmatrix} 3 & 0 & -2 \\ 0 & 4 & 0 \\ -2 & 0 & 3 \end{pmatrix}$ **23.** $\begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ **24.** $\begin{pmatrix} 2 & -3 & 4 \\ -3 & 2 & 0 \\ 4 & 0 & 2 \end{pmatrix}$

25. Find the C matrix which diagonalizes the matrix M of Problem 18. Observe that M is not symmetric, and C is not orthogonal (see Section 11). However, C does have an inverse; find C^{-1} and show that $\mathrm{C}^{-1}\mathrm{M}\mathrm{C} = \mathrm{D}$.

26. Repeat Problem 25 for Problem 19.

In Problems 27 to 30, rotate the given quadric surface to principal axes. What is the name of the surface? What is the shortest distance from the origin to the surface?

FEUE WHEP

27. $x^2 + y^2 - 5z^2 + 4xy = 15$

28. $7x^2 + 4y^2 + z^2 - 8xz = 36$

29. $3x^2 + 5y^2 - 3z^2 + 6yz = 54$

30. $7x^2 + 7y^2 + 7z^2 + 10xz - 24yz = 20$

31. Find the characteristic vibration frequencies of a system of masses and springs as in Figure 12.1 if the spring constants are k, $3k$, k.

32. Do Problem 31 if the spring constants are $6k$, $2k$, $3k$.

33. Prove the Caley-Hamilton theorem (Problem 11.60) for any matrix M for which $\mathrm{D} = \mathrm{C}^{-1}\mathrm{MC}$ is diagonal. See hints in Problem 11.60.

34. In problems 6.30 and 6.31, you found the matrices e^{A} and e^{C} (put $k = 1$) where A and C are the Pauli matrices from Problem 6.6. Now find the matrix $(\mathrm{A}+\mathrm{C})$ and its powers and so find the matrix $e^{\mathrm{A}+\mathrm{C}}$ to show that $e^{\mathrm{A}+\mathrm{C}} \neq e^{\mathrm{A}}e^{\mathrm{C}}$. See Problem 6.29.

35. Show that a square matrix A has an inverse if and only if $\lambda = 0$ is not an eigenvalue of A. *Hint:* Write the condition for A to have an inverse (Section 6), and the condition for A to have the eigenvalue $\lambda = 0$ (Section 11).

36. Write the three 3 by 3 matrices for 180° rotations about the x, y, z axes. Show that these three matrices commute (contrary to what we usually expect—see Problems 7.30 and 7.31). By writing the multiplication table, show that these three matrices with the unit matrix form a group. To which order 4 group is it isomorphic? *Hint:* See Problem 13.5.

37. Show that for a given irreducible representation of a group, the character of the class consisting of the identity is always the dimension of the irreducible representation. *Hint:* What is the trace of a unit n-by-n matrix?

38. For a cyclic group, show that every element is a class by itself. Show this also for an Abelian group.

FEUE WHEP

FÊUE WHEP

CHAPTER 4

Partial Differentiation

1. INTRODUCTION AND NOTATION

If $y = f(x)$, then dy/dx can be thought of either as the slope of the curve $y = f(x)$ or as the rate of change of y with respect to x. Rates occur frequently in physics; time rates such as velocity, acceleration, and rate of cooling of a hot body are obvious examples. There are also other rates: rate of change of volume of a gas with applied pressure, rate of decrease of the fuel in your automobile tank with distance traveled, and so on. Equations involving rates (differential equations) often need to be solved in applied problems. Derivatives are also used in finding maximum and minimum points of a curve and in finding the power series of a function. All these applications, and more, occur also when we consider a function of several variables.

Let z be a function of two variables x and y; we write $z = f(x, y)$. Just as we think of $y = f(x)$ as a curve in two dimensions, so it is useful to interpret $z = f(x, y)$ geometrically. If x, y, z are rectangular coordinates, then for each x, y the equation gives us a value of z, and so determines a point (x, y, z) in three dimensions. All the points satisfying the equation ordinarily form a surface in three-dimensional space (see Figure 1.1). (It might happen that an equation would not be satisfied by any real points, for example $x^2 + y^2 + z^2 = -1$, but we shall be interested in equations whose graphs are real surfaces.) Now suppose x is constant; think of a plane $x =$ const. intersecting the surface (see Figure 1.1). The points satisfying $z = f(x, y)$ *and* $x =$ const. then lie on a curve (the curve of intersection of the surface and the $x =$ const. plane; this is AB in Figure 1.1). We might want the slope, maximum and minimum points, etc., of this curve. Since z is a function of y (on this curve), we might write dz/dy for the slope. However, to show that z is actually a function of two variables x and y with one of them (x) temporarily a constant, we write $\partial z/\partial y$; we call $\partial z/\partial y$ the partial derivative of z with respect to y. Similarly, we can

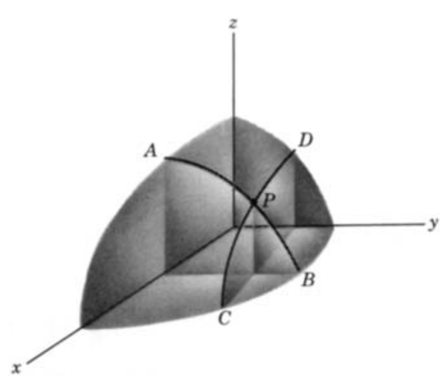


Figure 1.1

10431

hold y constant and find $\partial z/\partial x$, the partial derivative of z with respect to x. If these partial derivatives are differentiated further, we write

$$\frac{\partial}{\partial x}\frac{\partial z}{\partial x}=\frac{\partial^2 z}{\partial x^2},\qquad \frac{\partial}{\partial x}\frac{\partial z}{\partial y}=\frac{\partial^2 z}{\partial x\partial y},\qquad \frac{\partial}{\partial x}\frac{\partial^2 z}{\partial x\partial y}=\frac{\partial^3 z}{\partial x^2\partial y},\qquad \text{etc.}$$

Other notations are often useful. If $z=f(x,y)$, we may use z_x or f_x or f_1 for $\partial f/\partial x$, and corresponding notations for the higher derivatives.

Example. Given $z=f(x,y)=x^3y-e^{xy}$, then

$$\begin{aligned}
\frac{\partial f}{\partial x}&\equiv\frac{\partial z}{\partial x}\equiv f_x\equiv z_x\equiv f_1=3x^2y-ye^{xy},\\
\frac{\partial f}{\partial y}&\equiv\frac{\partial z}{\partial y}\equiv f_y\equiv z_y\equiv f_2=x^3-xe^{xy},\\
\frac{\partial^2 f}{\partial x\,\partial y}&\equiv\frac{\partial^2 z}{\partial x\,\partial y}\equiv f_{yx}\equiv z_{yx}\equiv f_{21}=3x^2-e^{xy}-xye^{xy},\\
\frac{\partial^2 f}{\partial x^2}&\equiv\frac{\partial^2 z}{\partial x^2}\equiv f_{xx}\equiv z_{xx}\equiv f_{11}=6xy-y^2e^{xy},\\
\frac{\partial^3 f}{\partial y^3}&\equiv\frac{\partial^3 z}{\partial y^3}\equiv f_{yyy}\equiv z_{yyy}\equiv f_{222}=-x^3e^{xy},\\
\frac{\partial^3 f}{\partial x^2\,\partial y}&\equiv\frac{\partial^3 z}{\partial x^2\,\partial y}\equiv f_{yxx}\equiv z_{yxx}\equiv f_{211}=6x-2ye^{xy}-xy^2e^{xy}.
\end{aligned}$$

We can also consider functions of more variables than two, although in this case it is not so easy to give a geometrical interpretation. For example, the temperature T of the air in a room might depend on the point (x,y,z) at which we measured it and on the time t; we would write $T=T(x,y,z,t)$. We could then find, say, $\partial T/\partial y$, meaning the rate at which T is changing with y for fixed x and z at one instant of time t.

A notation which is frequently used in applications (particularly thermodynamics) is $(\partial z/\partial x)_y$, meaning $\partial z/\partial x$ when z is expressed as a function of x and y. (Note two different uses of the subscript y; in the example above, f_y meant $\partial f/\partial y$. A subscript *on a partial derivative*, however, does *not* mean another derivative, but just indicates the variable being held constant in the indicated partial differentiation.) For example, let $z=x^2-y^2$. Then using polar coordinates r and θ, (recall that $x=r\cos\theta$, $y=r\sin\theta$, $x^2+y^2=r^2$), we can write z in several other ways. For each new expression let us find $\partial z/\partial r$.

$$\begin{aligned}
&z=x^2-y^2,\\
&z=r^2\cos^2\theta-r^2\sin^2\theta, &&\left(\frac{\partial z}{\partial r}\right)_\theta=2r(\cos^2\theta-\sin^2\theta),\\
&z=2x^2-x^2-y^2=2x^2-r^2, &&\left(\frac{\partial z}{\partial r}\right)_x=-2r,\\
&z=x^2+y^2-2y^2=r^2-2y^2, &&\left(\frac{\partial z}{\partial r}\right)_y=+2r.
\end{aligned}$$

These three expressions for $\partial z/\partial r$ have different values and are derivatives of three different functions, so we distinguish them as indicated by writing the second independent variable as a subscript. Note that we do *not* write $z(x,y)$ or $z(r,\theta)$; z is

FEUE WHEP

one variable, but it is equal to several *different* functions. Pure mathematics books usually avoid the subscript notation by writing, say, $z = f(r,\theta) = g(r,x) = h(r,y)$, etc.; then $(\partial z/\partial r)_\theta$ can be written as just $\partial f/\partial r$, and similarly

$$\left(\frac{\partial z}{\partial r}\right)_x = \frac{\partial g}{\partial r} \qquad \text{and} \qquad \left(\frac{\partial z}{\partial r}\right)_y = \frac{\partial h}{\partial r}.$$

However, this multiplicity of notation ($z = f = g = h$, etc.) would be inconvenient and confusing in applications where the letters have *physical* meanings. For example, in thermodynamics, we might need

$$\left(\frac{\partial T}{\partial p}\right)_v, \qquad \left(\frac{\partial T}{\partial v}\right)_s, \qquad \left(\frac{\partial T}{\partial p}\right)_u, \qquad \left(\frac{\partial T}{\partial s}\right)_p, \qquad \text{etc.,}$$

as well as many other similar partial derivatives. Now T means temperature (and the other letters similarly have physical meanings which must be recognized). If we wrote $T = A(p,v) = B(v,s) = C(p,u) = D(s,p)$ and similar formulas for the eight commonly used quantities in thermodynamics, each as functions of pairs from the other seven, we would not only have an unwieldy system, but the physical meaning of equations would be lost until we translated them back to standard letters. Thus the subscript notation is essential.

> The symbol $(\partial z/\partial r)_x$ is usually read "the partial of z with respect to r, with x held constant." However, the important point to understand is that the notation means that z has been written as a function of the variables r and x *only*, and then differentiated with respect to r.

A little experimenting with various functions $f(x,y)$ will probably convince you that $(\partial/\partial x)(\partial f/\partial y) = (\partial/\partial y)(\partial f/\partial x)$; this is usually (but not always) true in applied problems. It can be proved (see advanced calculus texts) that if the first and second order partial derivatives of f are continuous, then $\partial^2 f/\partial x\,\partial y$ and $\partial^2 f/\partial y\,\partial x$ *are* equal. In many applied problems, these conditions are met; for example, in thermodynamics they are normally assumed and are called the reciprocity relations.

▶ PROBLEMS, SECTION 1

1. If $u = x^2/(x^2+y^2)$, find $\partial u/\partial x$, $\partial u/\partial y$.

2. If $s = t^u$, find $\partial s/\partial t$, $\partial s/\partial u$.

3. If $z = \ln\sqrt{u^2+v^2+w^2}$, find $\partial z/\partial u$, $\partial z/\partial v$, $\partial z/\partial w$.

4. For $w = x^3 - y^3 - 2xy + 6$, find $\partial^2 w/\partial x^2$ and $\partial^2 w/\partial y^2$ at the points where $\partial w/\partial x = \partial w/\partial y = 0$.

5. For $w = 8x^4 + y^4 - 2xy^2$, find $\partial^2 w/\partial x^2$ and $\partial^2 w/\partial y^2$ at the points where $\partial w/\partial x = \partial w/\partial y = 0$.

6. For $u = e^x \cos y$,

(a) verify that $\partial^2 u/\partial x\partial y = \partial^2 u/\partial y\partial x$;

(b) verify that $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = 0$.

FEUE WHEP

If $z = x^2 + 2y^2$, $x = r\cos\theta$, $y = r\sin\theta$, find the following partial derivatives.

7. $\left(\frac{\partial z}{\partial x}\right)_y$ **8.** $\left(\frac{\partial z}{\partial x}\right)_r$ **9.** $\left(\frac{\partial z}{\partial x}\right)_\theta$ **10.** $\left(\frac{\partial z}{\partial y}\right)_x$ **11.** $\left(\frac{\partial z}{\partial y}\right)_r$ **12.** $\left(\frac{\partial z}{\partial y}\right)_\theta$

13. $\left(\frac{\partial z}{\partial \theta}\right)_x$ **14.** $\left(\frac{\partial z}{\partial \theta}\right)_y$ **15.** $\left(\frac{\partial z}{\partial \theta}\right)_r$ **16.** $\left(\frac{\partial z}{\partial r}\right)_\theta$ **17.** $\left(\frac{\partial z}{\partial r}\right)_x$ **18.** $\left(\frac{\partial z}{\partial r}\right)_y$

19. $\frac{\partial^2 z}{\partial r\,\partial y}$ **20.** $\frac{\partial^2 z}{\partial x\,\partial \theta}$ **21.** $\frac{\partial^2 z}{\partial y\,\partial \theta}$ **22.** $\frac{\partial^2 z}{\partial r\,\partial x}$ **23.** $\frac{\partial^2 z}{\partial r\,\partial \theta}$ **24.** $\frac{\partial^2 z}{\partial x\,\partial y}$

7′ to **24′**. Repeat Problems 7 to 24 if $z = r^2\tan^2\theta$.

2. POWER SERIES IN TWO VARIABLES

Just as in the one-variable case discussed in Chapter 1, the power series (about a given point) for a function of two variables is unique, and we may use any convenient method of finding it (see Chapter 1 for methods).

Example 1. Expand $f(x, y) = \sin x \cos y$ in a two-variable Maclaurin series. We write and multiply the series for $\sin x$ and $\cos y$. This gives

$$\sin x \cos y = \left(x - \frac{x^3}{3!} + \cdots\right)\left(1 - \frac{y^2}{2!} + \cdots\right) = x - \frac{x^3}{3!} - \frac{xy^2}{2!} + \cdots .$$

Example 2. Find the two-variable Maclaurin series for $\ln(1 + x - y)$. We replace x in equation (13.4) of Chapter 1 by $x - y$ to get

$$\begin{aligned}\ln(1 + x - y) &= (x - y) - (x - y)^2/2 + (x - y)^3/3 + \cdots\\ &= x - y - x^2/2 + xy - y^2/2 + x^3/3 - x^2y + xy^2 - y^3/3 + \cdots .\end{aligned}$$

The methods of Chapter 1, used as we have just shown, provide an easy way of obtaining the power series for many simple functions $f(x, y)$. However, it is also convenient, for theoretical purposes, to have formulas for the coefficients in the Taylor series or the Maclaurin series for $f(x, y)$; see, for example, Problem 8.2. Following a process similar to that used in Chapter 1, Section 12, we can find the coefficients of the power series for a function of two variables $f(x, y)$ (assuming that it can be expanded in a power series). To find the series expansion of $f(x, y)$ about the point (a, b) we write $f(x, y)$ as a series of powers of $(x - a)$ and $(y - b)$ and then differentiate this equation repeatedly as follows.

$$\begin{aligned}f(x, y) &= a_{00} + a_{10}(x - a) + a_{01}(y - b) + a_{20}(x - a)^2 + a_{11}(x - a)(y - b)\\ &\quad + a_{02}(y - b)^2 + a_{30}(x - a)^3 + a_{21}(x - a)^2(y - b)\\ &\quad + a_{12}(x - a)(y - b)^2 + a_{03}(y - b)^3 + \cdots .\end{aligned}$$

$$(2.1)\qquad \begin{aligned}f_x &= a_{10} + 2a_{20}(x - a) + a_{11}(y - b) + \cdots ,\\ f_y &= a_{01} + a_{11}(x - a) + 2a_{02}(y - b) + \cdots ,\\ f_{xx} &= 2a_{20} + \text{ terms containing } (x - a) \text{ and/or } (y - b),\\ f_{xy} &= a_{11} + \text{ terms containing } (x - a) \text{ and/or } (y - b).\end{aligned}$$

[We have written only a few derivatives to show the idea. You should be able to calculate others in the same way (Problem 7).] Now putting $x = a$, $y = b$ in (2.1),

FEUE WHEP